QUANTIFYING THE PRESENT

AND

PREDICTING THE PAST

Theory, Method, and Application of Archaeological Predictive Modeling

$10.75

QUANTIFYING THE PRESENT

AND

PREDICTING THE PAST

QUANTIFYING THE PRESENT AND PREDICTING THE PAST: THEORY, METHOD, AND APPLICATION OF ARCHAEOLOGICAL PREDICTIVE MODELING

edited by

W. James Judge and Lynne Sebastian

with contributions by

Jeffrey H. Altschul
James I. Ebert
C. Kincaid
Timothy A. Kohler
Kenneth L. Kvamme
Daniel W. Martin
Christopher L. Nagle
Martin R. Rose
Alston V. Thoms

technical editor

M. June-el Piper

for the
U.S. Department of the Interior
Bureau of Land Management
Denver, Colorado

December 1988

The contents of this publication do not necessarily reflect the views and policies of the Federal Government, U.S. Department of the Interior, or the Bureau of Land Management unless so designated by other authorized documents. Use or mention of trade names or commercial products is solely for the information and convenience of the reader and does not constitute endorsement or recommendation for use unless so designated by other authorized documents.

BLM/YA/PT-88/008+8100

Library of Congress Cataloging in Publication Data

Judge, W. James, and Lynne Sebastian, eds.
1988. Quantifying the Present and Predicting the Past: Theory, Method, and Application of Archaeological Predictive Modeling. U.S. Department of the Interior, Bureau of Land Management Service Center. Denver, Co. xx, 690 pp.

88-600552

For sale by the Superintendent of Documents.
U.S. Government Printing Office, Washington, D.C. 20402

LIST OF CONTRIBUTORS

Jeffrey H. Altschul (Ph.D., Brandeis University, 1981) President, Statistical Research (Consultants for the Biomedical, Environmental, and Social Sciences), 2500 North Pantano, Suite 218, Tucson, Arizona.

James I. Ebert (Ph.D., University of New Mexico, 1986) President, Ebert & Associates (Anthropology, Archaeology, Forensics and Environmental Sciences; Remote Sensing, Photogrammetric, and Geographic Information Systems Applications), 3100 Ninth St. NW, Albuquerque, New Mexico.

W. James Judge (Ph.D., University of New Mexico), Director, Fort Burgwin Research Center, P.O. Box 300, Ranchos de Taos, New Mexico.

Chris Kincaid, Archeologist, Glen Canyon National Recreation Area, P.O. Box 1570, Page, Arizona.

Timothy A. Kohler (Ph.D., University of Florida, 1978) Associate Professor, Department of Anthropology, Washington State University, Pullman, Washington.

Kenneth L. Kvamme (Ph.D., University of California, Santa Barbara, 1983) Associate Curator of Archaeology, Arizona State Museum, University of Arizona, Tucson, Arizona.

Daniel W. Martin, Archeologist, Bureau of Land Management, Service Center, Building 50, Denver Federal Center, Denver, Colorado.

Christopher L. Nagle (Ph.D., Brandeis University, 1984) Research Archaeologist, Department of Anthropology, Smithsonian Institution, Washington, D.C.

Martin R. Rose (Ph.D. Candidate, University of Arizona) Research Associate, Laboratory of Tree-Ring Research, University of Arizona, Tucson, Arizona.

Lynne Sebastian (Ph.D., University of New Mexico, 1988) Deputy State Historic Preservation Officer, Office of Cultural Affairs, 228 East Palace Avenue, Santa Fe, New Mexico.

Alston V. Thoms Senior Staff Archaeologist, Center for Northwest Anthropology, Washington State University, Pullman, Washington.

TABLE OF CONTENTS

Acknowledgments xvii

1. PREDICTING THE PAST: CORRELATION, EXPLANATION, AND THE USE OF ARCHAEOLOGICAL MODELS, *by Lynne Sebastian and W. James Judge* 1

   Models and Archaeology 1
   The Problem of Explanation 3
   History of the BLM Predictive Modeling Project 9
   The Production of This Volume 12
   The Structure of This Book 13
   Reference Cited 18

2. PREDICTIVE LOCATIONAL MODELING: HISTORY AND CURRENT PRACTICE, *by Timothy A. Kohler* 19

   The Wider Historical Debate: How and To What Extent Does The Natural Environment Influence Human Behavior? 22
   The Emergence of Settlement Pattern Studies in Archaeology 30
   The Era of Predictive Modeling 33
   Conclusions 52
   References Cited 53

3. MODELS AND THE MODELING PROCESS, *by Jeffrey A. Altschul* 61

   Types of Models 63
   The Model-Building Process 76
   References Cited 89

4. THE THEORETICAL BASIS OF ARCHAEOLOGICAL PREDICTIVE MODELING AND A CONSIDERATION OF APPROPRIATE DATA-COLLECTION METHODS, *by James I. Ebert and Timothy A. Kohler* 97

   Prediction, Models, and the Scientific Framework of Archaeology 101
   The Nature and Organization of Human Systems: Settlement, Mobility, and Technology 105
   Technological Strategies, Discard Behavior, and the Archaeological Record 116
   Natural Formation Processes and the Archaeological Record 123
   Ecosystems Variables and Archaeological Explanation and Modeling 128
   Distributional Archaeology 143
   Summary 158
   References Cited 162

5. AN OVERVIEW OF STATISTICAL METHOD AND THEORY FOR QUANTITATIVE MODEL BUILDING, *by Martin R. Rose and Jeffrey H. Altschul* 173

Modeling Site Location 175
Variables and Scales 182
Statistical Description and Inference in the Model-Building Process 200
Defining Site Classes 204
Modeling Techniques 212
Model Validation and Generalization 242
Conclusions 250
References Cited 252

6. COLLECTING NEW DATA FOR THE PURPOSE OF MODEL DEVELOPMENT, *by Jeffrey H. Altschul and Christopher L. Nagle* 257

Planning for Fieldwork 258
Survey Strategies 260
Data Collection in CRM Contexts 277
Data Processing 290
Conclusions 293
References Cited 294

7. USING EXISTING ARCHAEOLOGICAL SURVEY DATA FOR MODEL BUILDING, *by Kenneth L. Kvamme* 301

Use of Existing Data for Site-Location Models 302
Problems and Biases in Existing Site Survey Data 304
Procedures for Reducing Deficiencies and Biases in Existing Data 307
Evaluation of Site-Location Patterning and Model Building with Existing Data 313
Example Analysis 317
References Cited 322

8. DEVELOPMENT AND TESTING OF QUANTITATIVE MODELS, *by Kenneth L. Kvamme* 325

Variables Used in Locational Research 331
Assessing Patterns in Archaeological Locational Data 339
Application Comparison of Quantitative Locational Models 363
Combining Models for Locational Characteristics and Models for Location Only 378
Modeling Individual Site Types 381
Interpretation and Explanation of Data Patterns 386
Assessing Model Performance 389
Model Revision 415
References Cited 418

9. REMOTE SENSING IN ARCHAEOLOGICAL PROJECTION AND PREDICTION, *by James I. Ebert* 429

Fundamentals of Remote Sensing 430
Contemporary Applications of Remote Sensing to Archaeological Projection and Prediction 437

Potential Applications of Remote Sensing within the Explanatory Framework of Archaeological Modeling and Prediction 465
References Cited 485

10. GEOGRAPHIC INFORMATION SYSTEMS: TECHNICAL AIDS FOR DATA COLLECTION, ANALYSIS, AND DISPLAY, *by Kenneth L. Kvamme and Timothy A. Kohler* 493

Introduction 493
The Potential of Geographic Information Systems for Research, Development, and Application of Archaeological Site Location Models 496
Brief Overview of Three Common Geographic Information Systems 535
Model Display vs Model Building 539
Imaginary Session with a Generic GIS 540
References Cited 543

11. PREDICTIVE MODELING AND ITS RELATIONSHIP TO CULTURAL RESOURCE MANAGEMENT APPLICATIONS, *by C. Kincaid* 549

What Are Models About? 550
What Conditions Are Favorable for Modeling Projects? 551
What Kinds of Models Are There? When Do We Use Which Type? 554
How Can We Prepare for Model Development? 556
How Do We Plan a Model? 558
How Do We Apply Modeling in Cultural Resource Management? 560
How Can Models Be Evaluated? 564
Future Directions 567

12. AN APPRAISAL, *by W. James Judge and Daniel W. Martin* 571

Evaluation of Project Goals 572
An Appraisal of the Review Comments 573
The Issues Raised 574
Conclusions 579
References Cited 580

Appendix A SURVEY OF PREDICTIVE LOCATIONAL MODELS: EXAMPLES FROM THE LATE 1970s AND EARLY 1980s, *by Alston V. Thoms* 581

Tabulated Survey Results 583
Synopsis of Survey Results 585
Concluding Comments 636
References Cited 641

Index 647

LIST OF ILLUSTRATIONS

2.1 A suggested taxonomy for the different types of locational models that appear in the literature 36
3.1 The model-building process 77
4.1 The explanatory framework of archaeological science 103
4.2 The half-radius continuous pattern of exploitation of the landscape by foraging groups 110
4.3 The complete radius leapfrog pattern of landscape use 111
4.4 Parallel levels of ecosystems and settlement systems and their archaeological manifestations 129
4.5 Suggested effects of increasing intensification on the location of residential sites 136
4.6 Suggested effects of increasing intensification on the concentration and visibility of archaeological materials 137
4.7 Suggested effects of the spatial characteristics of critical environmental resources on locational modeling 140
4.8 Suggested effects of the temporal characteristics of critical environmental resources on locations of archaeological materials 141
4.9 Location of the Seedskadee Project 149

5.1 Diagram of a region 176
5.2 Relative frequency of E 177
5.3 Relative frequency of the composite result E 178
5.4 Venn diagram showing distribution of sites within piñon-juniper and desert shrub zones 179
5.5 Two probability functions 188
5.6 Constant density between X_1 and X_2 implies equal probability for the intervals I_1 and I_2, both of equal length 188
5.7 A highly centralized density function 189
5.8 Two densities differing in location 190
5.9 Histogram of discriminant function scores for Booth Mountain nonsite locations 230
5.10 Histogram of discriminant function scores for Booth Mountain site group 230
5.11 Histogram of canonical discriminant function scores for both nonsite and site locations, Booth Mountain 231
5.12 Posterior probability of site presence, Booth Mountain 249

6.1 Edge effect for survey units with different sizes and shapes 263
6.2 Bar graph of the frequency of quadrats by the number of sites, showing the positively skewed distribution 266

7.1 Relationship between crew spacing and survey rate 305
7.2 Errors in the locations of sites resulting from copying site locations from one map to another 307

7.3 Illustrations for bias correction procedures 310
7.4 GIS-generated images: locations of recorded Mesolithic sites and model of Mesolithic site location 319

8.1 End products of cultural resource modeling 328
8.2 Measurement of variables: slope as percent grade and aspect, local relief, and terrain texture 334
8.3 Measurement of variables: view angle 336
8.4 Models based only on locational or positional information 342
8.5 Uses of control data in empirical studies 348
8.6 Sampling practices 354
8.7 Robust classification models based on characteristics of locations 368
8.8 Simple mathematical models based on characteristics of locations 375
8.9 Combination of location-only and locational characteristics models 379
8.10 Mesolithic site-type locational model mapped through a GIS 385
8.11 Estimated site-class *p*-values for the initial nine-variable Glade Park site location model 391
8.12 Glade Park model performance curves 397
8.13 Sequential sampling design for a southern Arkansas study 408
8.14 Illustration of Bayes's Theorem and the effects of site and nonsite a priori probabilities 413

9.1 An eighteenth-century mill or industrial site discovered in Shenandoah National Park through remote-sensing-aided archaeological projection 441
9.2 Digitally derived reflectance values for sites vs nonsites in southwestern New Mexico taken from Landsat MSS computer compatible tapes 443
9.3 Sand ridge locations in Maryland with high probability of archaeological site occurrence 445
9.4 An ecologic/cover-type map of the National Petroleum Reserve in Alaska compiled with the aid of interpretation of Landsat color composite visual data 448
9.5 San Juan Basin ecologic/cover-type zones delineated through the interpretation of Landsat MSS visual data 452
9.6 A site-occurrence probability map of a Colorado study area compiled through the digital analysis of Landsat MSS data 458
9.7 A site-occurrence probability map derived through digital analysis of Landsat MSS data at the Naval Weapons Center, China Lake, California 459
9.8 A probability map showing the potential for the occurrence of archaeological sites from six cultural/temporal periods in Delaware 461
9.9 The location of an experiment in projecting archaeological site occurrence using simulated SPOT data in Bandelier National Monument, north-central New Mexico 462

9.10 A taxonomy of cultural and noncultural postdepositional processes 468
9.11 A portion of a map of geomorphological surface units at Chaco Culture National Historical Park derived using photointerpretation of 1:12,000 scale color aerial photos 471
9.12 Geomorphological surface units interpreted visually from a Landsat color composite print and 1:80,000 scale aerial photographs in and around the Seedskadee National Wildlife Refuge on the Green River, southwestern Wyoming 477
9.13 A computer-generated image showing a digital classification of cover-type units along Gallegos Wash in the San Juan Basin, northwestern New Mexico 483
9.14 A digitally derived environmental diversity index resulting from further computer analysis of the cover-type classification shown in Figure 9.13 484

10.1 Manual digitizing of contour lines through use of a cursor and digitizing tablet 502
10.2 Construction of a GIS 503
10.3 Characteristics of map information in vector-based and cell-based geographic information systems 505
10.4 Effects of cell size in cell-based geographic information systems 507
10.5 Steps that might be followed in the construction of an elevation surface 509
10.6 Elevation interpolation algorithms 510
10.7 Illustration of smoothing 512
10.8 Encoding of hydrologic data 513
10.9 Examples of various computational algorithms 514
10.10 Illustration of distance calculation techniques 515
10.11 GIS-generated surfaces: slope and aspect 522
10.12 GIS-generated surfaces: distance to nearest drainage and nearest second-order or greater drainage 524
10.13 GIS-generated surfaces: local relief and archaeological site sensitivity 526
10.14 Histograms of Euclidean distances to the nearest second-order or greater stream 530
10.15 Illustration of successive and cumulative selections made by a GIS 533

A.1 Map of North America showing approximate locations of predictive modeling projects discussed in the Appendix 582

LIST OF TABLES

2.1 Resources categorized by return rate and role in influencing site location 48
2.2 Example predictions 49

3.1 Types of objective predictive models of site locations 64
3.2 Bisti–Star Lake region site classes 70

4.1 Selected correlates of intensification 132

5.1 Record of simple results and r.f.*(E)* 177
5.2 Record of how the relative frequency r.f. *(E)* of the composite result *E* varies with the number *N* of trials 178
5.3. Means and standard deviations of variables for Booth Mountain nonsites, sites, and both groups combined 221
5.4. Covariance matrix for Booth Mountain nonsite group 223
5.5. Covariance matrix for Booth Mountain site group 223
5.6 Total covariance matrix based on all Booth Mountain cases, both nonsites and sites 225
5.7. Booth Mountain within-group covariance matrix 226
5.8. Booth Mountain within-group correlation matrix 226
5.9 Booth Mountain unstandardized discriminant function coefficients 227
5.10 Booth Mountain standardized discriminant function coefficients 228
5.11 Pooled within-group correlations between the canonical discriminant function and the discriminating variables 232
5.12 Classification functions for the Booth Mountain nonsite and site groups 234
5.13 Booth Mountain discriminant scores and probabilities of group membership 236
5.14 Stepwise output from Booth Mountain discriminant function analysis 239
5.15 Jackknife classification of locations from Booth Mountain 245

6.1 Skewness values and sample sizes by study tract 267
6.2 Probability of selecting a specified number of survey units in a particular zone under simple random sampling 271
6.3 Estimated density of prehistoric sites in project areas in and near the San Rafael Swell 278
6.4 Hypothetical four-level hierarchical field data file 285

7.1 Example of weights applied to data as a means of bias correction 312
7.2 Assessing model performance along several site-type and environmental categories 315
7.3 Descriptive statistics for the site location study of German Mesolithic sites 320

8.1 Descriptive statistics for Glade Park sites and nonsites 359
8.2 Classification performance of initial Glade Park models 362
8.3 Values for eight environmental variables 366
8.4 Classification performance rates of several site location modeling procedures 367
8.5 Comparison of classification performances of site location models 381
8.6 Illustration of changing cutoff p-values and their effects on classification accuracy 392
8.7 Original nine-variable Glade Park site location model, 36 jackknifed models, and final composite jackknifed model 398
8.8 Assessment of model classification results 401

9.1 Sites located by the NPR-A cultural resource surveys, 1977 and 1978 450
9.2 A 25 by 25 pixel matrix of simulated SPOT data over part of the Bandelier National Monument study area 464
9.3 Surface geomorphological units at Chaco Culture National Historical Park, northwestern New Mexico 472
9.4 Occurrence of known archaeological sites and materials at Chaco Culture National Historical Park within geomorphic surface units mapped with photointerpretive techniques 474

10.1 Selected training opportunities in geographic information systems 495
10.2 Selected journals and conference proceedings containing advanced discussions of geographic information systems 496
10.3 Statistical functions available in three commonly utilized systems 536
10.4 Data themes available for your GIS session 541

A.1 Summary of general characteristics of the selected predictive model projects 586
A.2 Summary of characteristics of data base and predictions from selected predictive model projects 592
A.3 Summary of environment and environmental variables used in selected predictive models 597
A.4 Summary of cultural identification and specificity of variables used in selected predictive models 600
A.5 Summary of general characteristics of field investigations conducted in conjunction with selected predictive model projects . . 604

ACKNOWLEDGMENTS

Any project of this magnitude could not be completed without the assistance of countless persons, including fellow professionals, clerical and administrative staff, and friends and family members. Personal acknowledgments of several of the contributors are presented at the end of the chapters. The following are persons and institutions to whom the editing team and the project advisory team owe unending gratitude.

The editing team would like to thank Rick Kirkness, Kathy Ives, Pam Dandrea and Kelly Lennick of the BLM Montana State Office in Billings, who were responsible for the typesetting and paste-up of this volume. Against all odds, and working with immutable deadlines, they have still managed to produce this publication. We would also like to thank the following Denver Federal Center personnel: Janet Poorman (BLM editor) and Margaret McGinnis (BLM copy editor) assisted with proofreading the galleys and checking the page proofs, Peter Doran designed the cover, and Carl Anderson, Mary Gaylord, Shirley L. McCulloch, and Carl Zulick worked on volume design and layout. Dan Martin and Mike Garratt had the unenviable task of compiling the index. Chuan Kun Ho produced the figures for the 1985 review draft of this volume, and Sam Tubiolo made corrections and drafted many new figures for the final publication. Jackie Rossignol assisted with the initial typing of the bibliographies for the review draft, and Janet D. Orcutt, who has an unusual combination of expertise in English, archaeology, and statistics, edited Chapter 5 for us. Finally, we especially wish to acknowledge Dan Martin, who as project coordinator, was ultimately responsible for bringing this project to a successful completion.

Many staff members of the BLM Service Center were extremely helpful in providing administrative support to this project. We would like to thank H. Robert Moore (current Director), Neil F. Morck (Director, 1984-1986), Delmar D. Vail (Director, 1983-1984), Irene Clapp (Contracting Officer), and the many bureau managers and supervisors who lent enthusiastic encouragement and support as well as ideas for this volume. Bob Ader provided administrative services, and Angie Sether and Janice Lopez provided clerical assistance. We would especially like to thank Mike Garratt, who read the entire work in its earliest stages and made many helpful comments on both specific and organizational problems.

This project was jointly conceived and compiled by members of the Project Advisory Team, who were involved throughout all stages of this effort:

Bureau of Land Management
- Daniel W. Martin (chair)
- Chris Kincaid
- Patricia Giorgi
- Ray Leicht

Rich Fike
Burton Williams
Forest Service
Jerry Wylie
Steve Sigstad
Society for American Archaeology
Keith Kintigh
SAA Regional Conferences
Cynthia Irwin-Williams and Don Fowler, co-chairpersons
C. Melvin Aikens
David Brose
Anne Marie Cantwell
David Frederickson
Garland Grabert
Thomas J. Green
Dale Henning
Cheryl Munson
Bert Salwen
R. Gwinn Vivian
J. Ned Woodall
William Workman

The following agencies, institutions, and individuals provided thoughtful and useful comments on the 1985 draft (copies of these comments may be requested from the Service Center):

Department of the Interior
Bureau of Land Management
National Park Service
Bureau of Reclamation
Office of Surface Mining
Bureau of Indian Affairs
Fish and Wildlife Service
Department of Agriculture
Soil Conservation Service
Forest Service
The Advisory Council on Historic Preservation
The National Council of State Historic Preservation Officers
State Historic Preservation Officers from
Arizona
Kentucky
Maryland
Ohio
New Mexico
South Carolina
South Dakota
Department of the Army, United States Army 5th Infantry Division

Alabama Historical Commission
Council of Alabama Archaeologists
American Petroleum Institute
Department of Anthropology, American University
Amoco Production Company
Arkansas Archaeological Survey
Department of Anthropology, University of California at Santa Barbara
Department of Anthropology, University of Delaware
Exxon Company USA
Division of Archives, History, and Records Management, Department of State, Tallahassee, Florida
Florida Archaeological Society
Kurt Steinhorst and Leo Flynn, University of Idaho
Laboratory of Anthropology, University of Idaho
Kentucky State Archaeologist, University of Kentucky
Michigan History Division, Department of State
Minneapolis Historical Society
Center for Archaeological Research, University of Mississippi
Department of Archives and History, Mississippi
New Mexico Archeological Council
Nickens and Associates
Powers Elevation Co.
Seneca Archeology Lab
Tetra Tech
Council of Virginia Archaeologists
Central Washington Archaeological Survey, Central Washington University
Wexpro Celsius
State Historical Society of Wisconsin

Chapter 1

PREDICTING THE PAST: CORRELATION, EXPLANATION, AND THE USE OF ARCHAEOLOGICAL MODELS

Lynne Sebastian and W. James Judge

MODELS AND ARCHAEOLOGY

One of the more interesting developments in the field of archaeology in the recent past is the emergence of predictive modeling as an integral component of the discipline. Within any developing and expanding field, one may expect some initial controversy that will, presumably, diminish as the techniques are tested, refined, and finally accepted. We are still very much in the initial stages of learning how to go about using predictive modeling in archaeology, and this book represents an effort by some of the leading experts in the field to present a comprehensive and detailed examination of this approach to understanding how people in the past used the landscape in which they lived.

There are probably as many definitions of the term *model* as there are scientific disciplines; several will be suggested in subsequent chapters of this book. We would like to offer a definition presented by David Clarke, who noted that models are "hypotheses or sets of hypotheses which simplify complex observations whilst offering a largely accurate predictive framework structuring these observations" (1968:32). There are two key aspects of this definition. The first is that models are selective abstractions, which of necessity omit a great deal of the complexity of the real world. Those aspects of the real world selected for inclusion in a model are assumed to be significant with respect to the interests and problem orientation of the person constructing the model. This is an important concept, since it indicates that there is no such thing as a truly objective model, be it inductively or deductively generated. Thus all models reflect, to a considerable degree, subjectivity on the part of the observer.

The second key aspect has to do with the predictive capability of models. Note that by this definition models have predictive content, and thus the term *predictive modeling* is somewhat redundant. We will employ this term here, however, since it has been widely accepted in archaeology.

This emphasis on the predictive aspects of models brings us to a more detailed examination of the concept of prediction itself, which the dictionary defines as "the ability to foretell on the basis of observation, experience, or scientific reason." One might even say that prediction is the essence of science because it allows us to formulate expectations about the future state of a system that are based on our knowledge of such systems or of similar ones (i.e., models). The point is that prediction is important, and that it is achieved scientifically through the generation of hypotheses that can be tested against the empirical record. Thus the *method* of prediction is essentially a deductive process, regardless of the form of generation of the model itself. Although the degree of formality might vary considerably, nearly all archaeological research today is based on a fundamentally deductive methodology.

Verification of formal predictive statements (hypotheses) through empirical testing against the archaeological record frequently involves techniques of sampling. In one sense all archaeology involves sampling, since we are never confronted with the complete record of past human behavior. Realizing this, archaeologists distinguish between relative degrees of sampling, as in "100 percent inventory" vs "sample survey." In this case, even though the results of both surveys are acknowledged to be samples, the latter term refers to a formally articulated, specific sampling strategy that guides the character of the inventory.

We mention sampling at this point because in the past formal sampling has frequently been confused with, and at times even identified with, predictive modeling; in the eyes of some, the implementation of a sampling design actually constitutes predictive modeling. Unfortunately, this confusion of sampling and predictive modeling has led to erroneous interpretations of the capabilities of the latter. Some researchers have even assumed that simply by adopting formal sampling techniques they would be able to predict archaeological site loci and thus satisfy legal compliance requirements without having to undertake expensive, 100 percent inventory surveys.

We would emphasize that sampling and predictive modeling are not the same thing and that formal sampling is neither required by predictive modeling nor limited to that approach. Sampling is simply one method of verifying testable hypotheses (albeit a very important one). In the strict sense—i.e., as a technique of data acquisition—formal sampling is no more (or less) related to or important to the modeling process than is 100 percent inventory survey.

One of the most unfortunate results of this confusion is that land-managing officials are at times led to believe that it is relatively easy to predict where all sites should be, and that by sampling a few of the predicted sites the archaeologists can do their jobs while saving themselves time and effort and saving the taxpayers a great deal of money. Realizing the distinction between sampling and prediction is a valuable first step in understanding how very complex the process of predictive modeling really is.

Both archaeologists and managers can and should be interested in refining attempts to model human behavior and in refining the sampling techniques used to

gather the data needed to verify such models. But neither models nor sampling should be viewed as a panacea destined to solve all the problems of management of archaeological resources and of compliance with existing legislation. This is a methodological fact of life that will be demonstrated repeatedly throughout this book.

THE PROBLEM OF EXPLANATION

Explanation in Archaeology

In the process of maturation, perhaps all scientific disciplines pass from a basically descriptive stage to a stage in which true explanation is attempted—a process of development that is sometimes painful and often divisive. The archaeological profession has been experiencing this transition for the past two decades, and the process has been both difficult and variably successful.

Twenty years ago, archaeology was a discipline in which most of the activity was directed toward describing the data that we recover. Since that time archaeologists have increasingly made conscious and consistent attempts to explain the changes in cultural process that were documented during the prior descriptive phase of archaeological research. It is obvious that such documentation must take place before explanation can be sought, but it is equally apparent that a discipline such as archaeology cannot remain at the descriptive level if it is to realize its full potential in contributing to scientific understanding.

Thus archaeologists who are undertaking the inventory and excavation of archaeological resources today are not simply concerned with accurately describing the artifacts and other data they find; they are equally concerned with placing those data in the context of explanation. That is, once they have determined what the things they recover are (or, more accurately, were) and how those things changed through time, they become interested in determining why such changes took place, in the *explanation* of such changes. In terms of the current jargon of our profession, we have progressed from dealing strictly with the archaeological context of the data to exploring their systemic context and finding means of linking the two realms.

As it has matured, archaeology has changed from a descriptive, documentary discipline to one that attempts to understand certain aspects of human behavior with reference to independent events and variables known to have occurred in the past. It is this attempt to understand human behavior that has given archaeology a new direction—a new sense of purpose, perhaps. Some would even say that this effectively legitimizes archaeology as a profession that is dependent in large part on public funding, but such a statement would evoke considerable argument among archaeologists themselves. In any case, most archaeologists would agree that we have progressed as a discipline, and that the new sense of purpose arising from

explanatory research emphases should be of concern to both archaeologists and land managers.

It is in the context of this transition from description to explanation that an important dichotomy apparent in this book arises. Those who read large sections of this book rather than using specific parts as a reference volume will soon notice that some authors focus on models that are deductively derived and attempt to predict how particular patterns of human land use will be reflected in the archaeological record while others are working with inductively derived models that identify and quantify relationships between archaeological site locations and environmental variables. The latter models, which we term *correlative*, are by far the more common in current modeling practice. It is our contention (and one that is shared by some but not all of the volume authors) that this emphasis on descriptive models will and should eventually be replaced by an emphasis on models that are derived from our understanding of human behavior and cultural systems, models with *explanatory* content.

The Value of Correlative Models

This call for explanation and explanatory models should not be taken as disparaging research that focuses on empirical analysis. Description, classification, and inductive generalizations are basic building blocks in any science. It should be clear from the sheer weight of information on correlative models in this volume and from the material presented in the management-oriented chapter (Chapter 11) that correlative models are informative and extremely valuable in many contexts.

In Chapter 11 Kincaid suggests that for some applications, simply knowing *where* sites are likely to be located relative to various environmental variables is sufficient. For large-scale planning purposes, for example, this level of knowledge about the distribution of archaeological resources may indeed be all that is needed for immediate purposes. But as suggested below, it may not be a wise use of resources to plan a research project solely to produce this level of information.

Several of the concepts introduced by Kvamme in the model applications chapter (Chapter 8)—those of *activity space* and *use intensity* in particular—make clear a second important contribution of correlative models. If a research project requires information about the *general* nature of human use of a landscape, correlative models provide invaluable data. It is both intuitively obvious and clear from the ethnographic record, for example, that human groups employing different subsistence strategies make use of their environments in very different ways. Their mobility patterns vary enormously, and the particular resources and proportions of those resources used are equally variable. In the archaeological record these differences will be reflected as differences in the scale of redundancy in distributions of cultural remains, that is, how big an area must be inspected before patterns in the archaeological record begin to repeat. Likewise, the nature and strength of correlations between cultural remains and features of the environment will be strongly affected

by differences in prehistoric resource selection. If we wish to monitor variability among human systems on the large scale, correlation models can provide a quantifiable and easily displayed measure of differences and similarities.

The Limitations of Correlative Models

Despite the utility of correlative models for planning purposes and for certain research applications, their general usefulness is limited for several reasons. The first is that no matter how carefully designed, methodologically sophisticated, and thoroughly tested a correlative model is, the end product is simply a series of statements about correlations between the occurrence of cultural remains and particular parameters or conjunctions of parameters of the *modern* environment. Correlation does not tell us anything about causality. We do not know, and cannot determine from the model, why this relationship between cultural materials and environmental factors exists. Worse yet, from an archaeological perspective, we do not know and cannot determine anything about the human system that created and deposited these cultural materials other than some very general notions about the distribution of their activities on the landscape.

The second limitation grows out of the first. Because correlative models are designed to tell us *where* sites are located (relative to various environmental variables) and not *why* they are located as they are with respect to those variables, even when they work exceedingly well, we do not know *why* they work. To the manager who only needs to know *where* sites are this may not immediately appear to be a major limitation. But if we do not know why a model works in one particular study area, we will not know whether we should expect it to work in the next valley or the next county or in a similar but distant environment. Thus correlative models are not truly predictive, but consist of projections of an observed pattern from a sample to the whole universe. When the focus of attention shifts to a new data universe, the process of projection must begin anew. As will be discussed in the next section, this lack of generalizability in correlative models should make this limitation of concern to managers as well as to the professional archaeologist.

The third limitation arises because correlative models require measurable, mappable data. For this reason, they depend heavily on environmental factors to provide their independent variables, and because of this they are most successful when applied to societies whose movements, group size, and activities are highly regulated by aspects of their environment—generally hunters and gatherers. With a shift from food collection to food production, human societies enter into a different kind of relationship with their environment (characterized by Kohler in Chapter 4 as one of increasing intensification). This does not mean that settlement locations of formative level societies cannot be modeled or that they are unresponsive to environmental factors. But the relationship with environmental factors is probably more indirect and is certainly more complex and interactive. Additionally, with increasing sedentism, social and political factors come to have an increasing impact

on the distribution of activities and thus of sites, lessening the correlations with strictly environmental variables.

Finally, because human groups with different subsistence orientations and different levels of technology use the landscape in very different ways, correlation models based on environmental variables are difficult to build for areas that have been occupied over a long period of time. In the American Southwest, for example, where the same area may have been used by Paleoindian, Archaic, Puebloan, and Athabaskan groups, a single correlation model of the relationships between cultural resources and environmental variables would be of very limited value. In such cases, an entire series of separately derived and tested models might be necessary, one for each major adaptation type.

The Value of Explanatory Models

The discussion above of the transition to explanation in modern American archaeology makes clear the importance of explanatory models to the archaeological profession and suggests that explanatory models are central to whatever value archaeology has for society as a whole. As anthropologists, we are interested in human behavior, in cultural variability and similarity, in cultural stability and change, in the adaptation of humans as cultural beings to their natural and social environments. As social scientists, we have an obligation to add to the store of human knowledge about humanity, and as archaeologists we have a unique opportunity to contribute knowledge about the long-term history of humankind, about adaptational successes and failures, and about the evolution of the complex social, political, and economic systems that order and dominate our lives.

If the value of explanatory models to archaeologists is clear, the value of these models to landholding agencies and to individuals involved in the field of cultural resource management is far less obvious. Because correlative models are relatively straightforward to develop and because simple environmental variables are relatively easy to measure, these models are viewed as cost-effective and objective. And in the short run they often provide the kinds of information needed. This has sometimes led to a perception on the part of managers that explanatory models are an unnecessary luxury. There are at least two reasons, however, why such models may, in the long run, prove to be critical to the very people who now question their utility or at least their cost-effectiveness.

The first reason has to do with the lack of generalizability for correlative models that was discussed above. If we do not know why a model works in one study area, we have no way of knowing *whether* it will work in a new study area; however much we may believe or expect that it will work, we cannot *know*. In order for a cultural resource manager to use information derived from models, even for the most general planning purposes, he or she must *know* that the model works within specified levels of confidence and precision. With correlative models, therefore, the process of model development, testing, refinement, and retesting can never be

short-cut: every new situation will require the development and verification of a new model.

With explanatory models, on the other hand, eventually we can hope to be able to offer general models that can be demonstrated to be applicable in any situation characterized by a specified set of cultural system and ecosystem variables. The key word here is, of course, "eventually"; as noted in the next section, explanatory models are extremely complex and difficult to build, and it may be a long while before we can be consistently successful in doing so. But that does not alter the potential value to resource managers of such powerful and truly generalizable models.

The second reason why explanatory models are potentially of great value in a management context has to do with the basic foundation of cultural resource management as it was envisioned in the National Historic Preservation Act (NHPA). One of the more colorful senior members of the American archaeological community admonishes his students not to lose sight of their major research objectives and become bogged down in trivia by reminding them that "It's hard to remember that you started out to drain the swamp when you're up to your [anatomical reference deleted] in alligators." Cultural resource management (CRM), especially as it is practiced in large land-managing agencies, tends to have the same problem. Sometimes we become so bogged down in the minutia of finding sites and protecting sites and mitigating impacts to sites that we lose track of the reason why these things called "sites" have any importance, any claim to protection under the law.

A great deal of time and energy is devoted to compliance with Section 106 of the NHPA, the section that mandates consideration of the impacts of federal undertakings on cultural resources and avoidance or mitigation of those impacts where possible. Sometimes this attention to Section 106 causes us to lose track of the requirements of Section 110, which charges federal agencies with the larger task of locating, inventorying, and nominating to the National Register of Historic Places the eligible properties under their control and instructs them to take care that these properties are not "inadvertently transferred, sold, demolished, substantially altered, or allowed to deteriorate significantly." In management terms, so much energy is going into the support program that the primary program gets slighted.

Probably the most commonly cited criterion for claiming National Register eligibility for a prehistoric site is that it has "yielded, or may be likely to yield, information important in prehistory or history" (36 CFR 60.4). It is their *information content* rather than any intrinsic value that gives archaeological sites significance and thus a legal right to protection, and it is because of this information content that the landholding agencies have been given a mandate to manage these resources.

It is the long-range goals of Section 110 compliance that can most benefit from the kind of understanding of the archaeological record that could be gained from explanatory models. For most archaeological sites discovered during the course of CRM-funded surveys, the survey recording and analysis will constitute the only scientific attention ever accorded to those sites. We would suggest, therefore, that

by calling for archaeological models that emphasize explanation rather than correlation, managers would maximize their return on realizing the information potential of the sites under their jurisdiction and find themselves in a better position to fulfill their responsibilities under Section 110. While correlation models might eventually become powerful and sophisticated enough to meet some of the requirements of Section 106 compliance, explanatory models could, in the long run, come much closer to meeting the need of compliance with Section 110.

The Limitations of Explanatory Models

The limitations of explanatory models are discussed by Altschul in Chapter 3, but his evaluation of the problems can be summed up in one short sentence: explanatory models are *extremely* difficult to create and validate. The length of the method and theory chapter (Chapter 4) and the complexity of the arguments presented therein by Ebert and Kohler make clear the difficulty of identifying the linkages and warranting the arguments in a model that is based in anthropological theory. The length of the model applications chapter (Chapter 8) and the complexity of the techniques discussed by Kvamme make it clear that currently correlative models are far ahead of explanatory models in methodological sophistication and mathematical expression.

The other serious limitation of explanatory models is one that is common to all attempts at explanation in archaeology. It has to do with assigning meaning to what we find in the archaeological record. In building an explanatory model we use information derived from the systemic context—often from ethnographic or ethnoarchaeological research, but sometimes from geography, ecology, or other fields—to generate hypotheses about the archaeological context. If we build these hypotheses into models and test them against the archaeological record and find that the results tend to confirm the model, then we assign meaning to the archaeological remains based on our interpretations of the systemic context.

The danger here is that our understanding of the systemic context will be incorrect. If we say that finding x in the archaeological record will mean that y happened in the systemic context, and if our ideas about y are wrong, then no matter what we find in the archaeological record our interpretations will be flawed. For example, until the late 1960s most archaeologists believed that hunters and gatherers lead an extremely difficult and precarious existence, teetering constantly on the brink of starvation and devoting every waking hour to the quest for food. Given such a perspective it seemed obvious that any hunter-gatherer group that had the opportunity to do so would immediately adopt agriculture, which was viewed as an easier and more secure way of life. Subsequent research demonstrated that hunting and gathering is, in fact, a rather stable and secure means of making a living and that agriculture is, in fact, both a more laborious and (in many environments) a less secure subsistence strategy. Most of the early archaeological research on the origins of agriculture was based on these incorrect notions about the

systemic context of hunting and gathering, and the results were, therefore, wrong or at least inadequate.

Although this danger of being fundamentally wrong is certainly an important limitation of explanatory models, it is also in a sense an indication of progress. As long as archaeologists concentrated solely on description and documentation it was nearly impossible for them to be wrong in any but trivial matters. But when they took the major step of attempting true explanation, they had to accept the risk of being profoundly wrong along with the rewards of gaining knowledge. The same relationship exists between correlative and explanatory models. Although there may be arguments about how to test for correlation or how to measure the strength of a correlation or assign confidence limits to it, once those are resolved the only question that remains is whether a correlation exists or not. With explanatory models the risks of being very wrong are much higher, but the potential gains in knowledge are correspondingly increased.

In the final analysis, we would suggest, a willingness to accept the risk of being wrong is one of the requirements of science. Scientific explanation consists of theories, statements about the way that we believe the world operates. An individual scientist offers an explanation that he or she believes accounts for as much variability in the phenomenon under study as possible. Subsequently this scientist and others test this explanation against data concerning the phenomenon, and the explanation is refined and revised to cover yet more of the variability. Empirical generalizations concerning the data can serve as one source of explanatory hypotheses, but those hypotheses cannot subsequently be tested against the same data. And empirical generalizations based on the archaeological record can never generate explanations of human behavior. We would argue that while correlative models are valuable in several contexts and explanatory models have several serious limitations, the ultimate goal of archaeological modeling, whether carried out for research purposes or to meet management needs, should be explanation.

HISTORY OF THE BLM PREDICTIVE MODELING PROJECT

In May 1983 a group of Bureau of Land Management (BLM) state archaeologists and Forest Service regional archaeologists from the Rocky Mountain states were meeting in Salt Lake City as part of a multistate task force designing procedures to deal with oil and gas development on public lands. During the course of these meetings, a number of informal discussions took place about the potential and problems of predictive modeling. It soon became clear that this was a subject of both great interest and great concern to the task force members, and a decision was made to begin a group project to study the ramifications and requirements of predictive modeling and to coordinate modeling efforts throughout the Mountain West.

As it happened, the Colorado State Office and Service Center of the BLM had recently initiated a predictive modeling study project, and with the support and cooperation of many people in the management hierarchy of the BLM, the newly

organized group of state and regional archaeologists was able to secure permission in September of 1983 to expand the scope of this already approved project to encompass an in-depth, state-of-the-art study of predictive modeling in archaeology. All those who had been at the task force meetings recognized that such a study was necessary if the problems encountered as a result of previous uses of predictive modeling in resource management contexts were to be avoided. This volume is the first product of the BLM Cultural Resource Predictive Modeling Project, but it is not the only product being planned. A training program and a technical assistance service for field personnel are planned, along with a set of demonstration models, which will be developed in future phases of this project.

In their proposal to expand the predictive modeling study to make it as comprehensive as possible, the Project Advisory Team (PAT; that is, the BLM and Forest Service archaeologists) pointed out that several predictive modeling attempts that had recently been carried out in management contexts had been highly controversial and of limited utility. They went on to add that since knowledge about this topic was limited among cultural resource professionals—both within the government and outside it—the lack of standards, guidelines, and procedures was hindering effective and efficient use of modeling for resource management.

The specific failings of past modeling efforts that they noted included failure to address management needs, lack of specificity, poor use of existing data, ineffective or biased sampling designs, inappropriate statistical analysis techniques, failure to collect inventory data suitable for the development of a predictive model, development of models using nonreplicable techniques, lack of comparability of and inappropriate use of environmental variables, lack of phasing to allow for model testing and refinement, and failure to use such technical aids as remote sensing and geographic information systems to streamline model development.

The stated goals of the expanded predictive modeling project were

1. to evaluate trends in the development of predictive modeling critically, using knowledge gained through past research;

2. to explore the feasibility and practicality of predictive modeling for meeting management objectives;

3. to analyze and define the components of the model-building process, particularly with respect to cultural resource management;

4. to develop a set of standards for the archaeological and environmental data to be used in modeling efforts; and

5. to provide BLM field offices with information on data collection for modeling purposes and statistical manipulations of those data.

The most important step in meeting these goals would be to contract with a team of outside consultants—archaeologists with national reputations in the field of predictive modeling—to produce a comprehensive, publishable report on this

topic. In addition, this project would have considerable input from BLM personnel, from a volunteer advisory group consisting of archaeologists for other federal agencies and individuals from State Historic Preservation Offices and the National Advisory Council on Historic Preservation, and from the professional archaeological community, including private contractors, representatives of professional organizations, and personnel from universities and museums. These individuals are named in the *Acknowledgments* at the front of this book.

To ensure that the profession at large would have the opportunity for a high level of input, several steps were taken. Once the expansion of the predictive modeling project had been approved, the PAT met at the Nevada State Office in Reno to determine how to organize and implement the project. As part of this meeting, the PAT met with representatives of the Society for American Archaeology (SAA) in an effort to secure society input and support for this project from its inception. The project team also corresponded with the society's president and executive committee, outlining the goals of the project and requesting suggestions for potential contractors and comments on the initial chapter outlines for the proposed book. In addition, members of the team met with regional representatives of the SAA to discuss the project and secure input, and the Procurement and Personnel Committee of the PAT held an open meeting for potential contractors and other interested persons at the 1984 annual meetings of the SAA in Portland, Oregon.

From the beginning of the project the BLM's Washington office provided normal intra-agency coordination among Washington, D.C., agencies. The PAT provided project briefings in Washington for top-level management and for senior-level agency archaeological program heads. Useful project direction was offered by these individuals, and most agreed to organize and provide a formal review of the initial draft document by their respective agencies.

Preliminary chapter outlines for the proposed predictive modeling book were prepared at the November 1983 meeting in Reno. Once these had been reviewed by the various advisory groups, final outlines were prepared, and requests for proposals were sent to potential contractors suggested by the various advisory groups and by members of the PAT. Those who wished to bid on one or more chapters responded with proposals that included detailed revised outlines for the chapters of interest. Successful bidders were selected on the basis of separate cost and technical proposals, with technical merit being more important than price. Since quality performance was considered vital to a successful project, the government reserved the right to award a contract on other than the lowest-price basis if a higher-priced proposal was rated higher in quality. The revised outlines submitted by the successful bidders were once again circulated to the advisory groups for comment, and then in August of 1984 the entire book-production team—authors, editors, and PAT—met in Denver for a prework conference.

THE PRODUCTION OF THIS VOLUME

At the prework meeting the authors and editors were given a cram course on the history and goals of this project, and then we attempted, in the course of several strenuous but exhilarating days, to give structure and coherence to this exercise in authorship by committee. We dealt ruthlessly with redundancies, struggled with what proved to be an insurmountable dichotomy among the authors in their view of the very nature of predictive modeling, and shifted the content and order of the chapters so many times that everyone (except the technical editor, who was keeping score) lost track of the "new" order by the third day.

One of the most difficult tasks of those days in Denver was to get a group of largely academic- and contract-oriented archaeologists to think in terms of management issues. Indeed, the very phrase "management concerns" produced mock groans by the end of the first day. We *did* gradually become more aware of the whole gamut of problems implied in the concept of management concerns, but it also became apparent to everyone that in writing and editing this book we could only do what we knew best—produce a book about predictive modeling; the real grappling with management concerns would have to be done by those who understood them best—the federal archaeologists of the PAT. At that point Dan Martin and Chris Kincaid, charter members of the PAT, agreed reluctantly to write the management issues chapter of the book with heavy input from the other team members; subsequently Burt Williams bowed to similar pressure and "volunteered" to be a coauthor on this chapter. By the end of the Denver meeting we had developed a final outline for the book and for each of the chapters, and the authors' difficulties began.

Between August of 1984 and January of 1985 most of the material in Chapters 2-10 of this book was written—an impressive feat given that all of the authors had simultaneous major commitments to teaching or to other contracts and writing responsibilities. In February of 1985, after we had a chance to at least skim most of the manuscripts, the editors and the PAT met to discuss the "product" and to make various editorial decisions. It was again clear that the main body of this book was not as management-oriented as the team members had hoped, but it was also clear that the manuscripts before us were the raw material of an invaluable resource volume—containing comprehensive, up-to-date treatments of the theoretical, methodological, and technical issues facing those who attempt to do archaeological predictive modeling. And again, this meant that the burden of meeting "management concerns" was going to lie wholly on the PAT members who were writing the management issues chapter. After this meeting, the editors' difficulties began.

In a slow, collaborative process between editors and authors (and taking the written comments of the PAT closely into account) we gradually shaped the individual manuscripts into the chapters of a generally unified book. As noted below, we made no effort to impose an artificial consistency of viewpoint on these authors. Archaeological predictive modeling is a field in which no consensus has emerged: that is one of the main points that is demonstrated in this book. When

the authors and editors had reached agreement on draft chapters, the book was sent out for a detailed and extensive peer review in October of 1985.

The reviewing agencies and organizations are also listed in the *Acknowledgments*. The review comments were compiled by the PAT and the volume editors, who carefully considered all comments and then summarized them by areas of concern. Minor questions or comments were handled by the editors; more substantial comments were forwarded to the authors, who responded in whatever way seemed appropriate and incorporated changes based on points raised by the reviewers into their various chapters. The results of the review are discussed in Chapter 12.

It was at this point that we hit the only major snag in the whole process of producing this volume. The Washington office of the BLM was not satisfied with the management concerns chapter and did not release it for review along with the rest of the book. Through a very long process of discussions between the PAT and the Washington office, it eventually became clear that Chapter 11 would have to be completely rewritten. Chris Kincaid once again accepted this task, and in 1988 she produced a draft of the chapter as it appears in this book. Chapter 11 and Chapter 12, the summary by Judge and Martin, were sent out for comment to a smaller corpus of reviewers selected from the large number of people who reviewed Chapters 1-10.

We have included this detailed discussion of the history of the BLM predictive modeling project and of this book because we, as editors, feel that this volume represents the culmination of a remarkable cooperative effort—something that *we* can say because the credit for those noteworthy aspects of this project lies with others. The determination and far-sightedness of the PAT members who conceived the notion of a large-scale, comprehensive, and high-quality effort and then guided, coaxed, and coerced the project into becoming a reality were certainly remarkable and commendable. Special merit accrues to Dan Martin and Chris Kincaid, who kept the project going during the long Chapter 11 delay and who wrote and rewrote the new Chapters 11 and 12 to solve the problems.

Finally, this book represents a remarkable degree of involvement and cooperation on the part of many people from all sectors of the archaeological profession. This has certainly contributed substantially to the quality of the book, but equally important, this level of cooperation seems to us to indicate that the often decried isolationism of academic, federal, and contract archaeologists may, like reports of Mark Twain's demise, have been greatly exaggerated.

THE STRUCTURE OF THIS BOOK

General Orientation

It will probably be helpful to the general reader to know four things about the overall orientation of this book at the outset. The first of these is that this is a book

about modeling in the context of prehistoric archaeology. While many of the principles suggested and techniques used would undoubtedly be of use to archaeologists studying classical and historical societies, particular problems and concerns of those scholars and techniques that would be especially helpful to them are touched on only in passing in this book. This orientation is a reflection of the background and experience of the authors and editors, and it is also a result of most of the extant predictive modeling studies having been concerned with prehistoric cultural remains.

The second thing, while we are on the subject of the intended audience for this volume, is that we have tried to maintain a balance between materials that would be of most interest to landholding agency managers and federal and state archaeologists and material that would be of interest to the archaeological profession in general. Certain chapters, such as the method and theory discussion by Ebert and Kohler in Chapter 4, will certainly be of greatest interest to professional archaeologists, while other chapters, such as the management perspectives chapter by Kincaid (Chapter 11) will be of greatest interest to managers. Still other chapters, such as the statistics discussion by Altschul and Rose (Chapter 5), will probably be viewed by readers of both persuasions as a resource document to be consulted as needed. The result of this effort to balance the book among somewhat disparate audiences is that nearly all readers will find some parts of the book more interesting than others. We have attempted, through our discussion below of the subjects covered in each chapter, through frequent cross-referencing, and through the production of a relatively detailed index, to enable the reader to identify quickly those subjects and discussions that are likely to be of interest to him or her.

The third thing to be noted is that even though some of the volume authors are strongly committed to the necessity for constructing explanatory models with major deductively derived components (see especially Ebert and Kohler in Chapter 4), by far the largest part of the book consists of information on correlative models derived largely or wholly through inductive means. These conflicting conceptions of the proper nature and direction of predictive modeling in archaeology are clear throughout the book; there was some discussion about the advisability of attempting to impose an editorial "synthesis" on the two camps of authors to create a theoretically and methodologically unified book, but we felt that this was artificial and premature. The division that is apparent in this book between those who are building sophisticated and fascinating correlative models and those who insist that archaeology is explanation or it is nothing is a reflection of the state of predictive modeling in American archaeology today. We felt that if this book was to be a fair summary of "the state of the art," the unresolved theoretical conflicts as well as the exciting technological and methodological advances should be explored. We have offered our own ideas on explanation and correlation in archaeological models in a previous section of this chapter, but we tried not to impose those ideas on the authors during the editing process.

The final point that we should raise about the general orientation of this book is that it is heavily biased toward models for hunter-gatherer societies. It was not

planned that way, and we tried to decrease this bias after the first draft of the book was finished. But we found that it was not that simple. In large part this emphasis on hunter-gatherers is a reflection of the emphasis on correlative models. As was noted in the discussion of the limitations of correlative models above, these models are most successful when applied to societies with a food-collecting subsistence base and relatively simple and fluid forms of social and demographic organization. In addition, this emphasis on hunter-gatherers seems to be a result of the interests of many of the researchers carrying out archaeological modeling projects today, so in this way the book is again a reflection of current developments in the field. We see this lack of modeling interest in middle-level or formative societies as well as historical societies as unfortunate, however, and would like to think that an increased interest in this topic will be one of the trends in future modeling projects.

A Preview of Coming Attractions

The main body of this book contains information that can roughly be divided into four topics. Chapters 2 through 4 present general discussions related to the modeling process. In Chapter 2 Kohler first reviews the intellectual history of what we today call predictive modeling, tracing the changing views of the relationship between human societies and their environment through time. He then discusses the contributions of the culture ecologists and especially that of Julian Steward to our thinking about this relationship. Finally, he describes the growing interest in predictive modeling in recent years and suggests a set of general criteria for evaluating models—generalizability, simplicity, internal consistency, precision, and falsifiability—using a group of example modeling projects to illustrate these concepts.

In Chapter 3 Altschul discusses models in general and the process of modeling. He suggests a typology of predictive models based on the spatial referent of the model and provides archaeological examples of the various types. He also discusses the methodological pitfalls of the various types of models and their strengths and weaknesses. Finally, he provides an overview of the model-building process, touching on data collection, synthesis, and evaluation; selection of independent and dependent variables; and model testing and refinement. All of these topics are addressed in detail in Chapters 5 through 8.

Chapter 4, by Ebert and Kohler, deals not with modeling as such, but with the theoretical and methodological considerations that must underlie all modeling efforts if the resultant models are to be faithful replications of human systems. Although the material presented is sometimes difficult, the concepts under discussion are, in the long run, just as critical to the success of modeling efforts as are questions of data collection or statistical manipulation. The authors discuss the organization of human systems and the implication of various organizational principles for the nature of the archaeological record produced. They also consider the relationship between human systems and the ecological systems of which they are a

part. Finally, they discuss the archaeological record itself—the way it is formed and the processes that affect it after the cultural materials are deposited—and offer suggestions about the implications of these formation and transformation processes for archaeology in general and for predictive modeling in particular.

Chapters 5 through 8 cover the details of the modeling process presented in overview in Chapter 3. In Chapter 5 Altschul and Rose discuss statistical approaches to modeling, particularly the theoretical and methodological considerations that must be taken into account in the course of building quantitative models. This chapter is not a cookbook of statistical techniques, but rather presents information on the general types of quantitative models. They discuss techniques of prediction and classification, emphasizing the strengths, limitations, and underlying assumptions of each, and describe various procedures for verifying the resultant models and generalizing from them.

Chapter 6, by Altschul and Nagle, covers the strategies and techniques involved in collecting new data for use in model development. The important and complex topic of sampling and the attendant problems of unit size and shape, sample size and means of selection, and techniques of parameter estimation are covered in detail. The authors also present a valuable discussion of the particular problems that arise when data must be collected within the constraints of cultural resource management surveys, where the survey universe and often the survey intensity are prescribed on the basis of considerations that have nothing to do with modeling requirements or research needs. Finally, they discuss various considerations of data recording, especially those imposed by "no collection" surveys.

In Chapter 7 Kvamme discusses the use of already collected data for model development, a topic of considerable importance given the quantity of existing data and the cost of data collection. As the author points out, the major problem with using existing data is that they very often are biased, and usually the type or types of biases present in the data base are unknown. He discusses the most common types of bias and suggests the effects that such biases will have on models developed using these data. He then offers a series of procedures for reducing deficiencies and minimizing the effects of biases. Finally, he describes ways of evaluating models built with existing data and means of determining what additional data must be collected in order to create a satisfactory model.

In Chapter 8 Kvamme goes on to discuss the actual steps in model building, beginning with the selection of variables and describing in detail various quantitative techniques for pattern recognition and assessment. He then considers the difficult problem of assessing model performance, discussing various means for measuring accuracy rates and assigning confidence limits to model results and providing a comparative analysis of several kinds of quantitative models.

Chapters 9 and 10 present information on types of technical aids that are available to assist researchers in the development of predictive models. In Chapter 9 Ebert summarizes the field of remote sensing, describing the devices used, the kinds of data that can be derived, and the types of analytical procedures commonly

applied to them. He then discusses the general potential of remote sensor data for predictive modeling applications and describes and evaluates several archaeological modeling projects that have involved the use of such data.

In Chapter 10 Kvamme and Kohler discuss a very exciting and relatively new technological aid, the Geographic Information System (GIS). A GIS comprises a set of computer programs, the hardware on which the programs run, and a spatially organized data base. In a GIS, data are derived from maps and similar sources of information on spatial relationships, and these data are stored not sequentially, as they are in most data base management applications, but in a form that retains the organizational information of the original data as well as the actual values of the variables. The applications of GIS discussed by Kvamme and Kohler make it clear that the potential of these systems for aiding in the predictive modeling process is enormous.

Finally, Chapter 11 is concerned with the federal management perspective on archaeological predictive modeling. The chapter is organized around a series of commonly asked questions, e.g., "What kinds of models are there? When do we use which type?" Kincaid summarizes relevant conclusions reached by the various authors and describes the potential usefulness of models for such central tasks of CRM as inventory, evaluation, resource protection, and planning.

In Chapter 12 Judge and Martin offer an appraisal both of the relative success or failure of the project in meeting the goals set for it originally and of the massive review process to which the draft manuscript was subjected. They then suggest several major issues raised in the course of this volume that they feel should be central questions in future modeling efforts.

The final section of this volume is an appendix compiled by Thoms, which presents an annotated review and assessment of a number of important and representative archaeological predictive modeling projects that have been carried out in recent years. The purpose of this appendix is to provide additional information on the kinds of projects that have been done, on the types of data that have been generated, and on the successes and pitfalls of such projects in the past.

We hope that this book will become a major reference volume for the archaeological profession as a whole as well as filling its original role in providing comprehensive, up-to-date information on topics related to predictive modeling for federal archaeologists and land-use managers. We feel that the blend of information offered here on modeling concepts, mathematical and statistical techniques, technical aids (such as remote sensing and GIS), and concerns about the relationship between modeling and archaeological method and theory will go a long way toward meeting the needs of researchers who are interested in this form of data analysis and interpretation and who wish to construct informed, sophisticated models.

In addition to the more general thanks expressed in the volume acknowledgments, we would like especially to thank two people. First we wish to express our appreciation to Dan Martin, whose gentle persistence, unfailing good humor, and determination kept the project afloat through delays, disasters, and aggravations and made this publication a reality. But most of all, we want to thank June-el Piper, the technical editor of this volume, who typed, edited, formatted, printed, read proofs, helped with indexing, cajoled authors, soothes savage beasts, and kept track of enough details to cause brown-out in lesser mortals.

REFERENCE CITED

Clarke, David L.
1968 *Analytical Archaeology*. Methuen, London.

Chapter 2

PREDICTIVE LOCATIONAL MODELING: HISTORY AND CURRENT PRACTICE

Timothy A. Kohler

In a volume primarily devoted to predicting locations of archaeological materials on the basis of factors in the natural environment, it seems important to spend a little time examining the anthropological underpinnings for such endeavors. In the first part of this chapter, relevant portions of the history of anthropological thought up to the 1940s are reviewed briefly and the contributions of Julian Steward are discussed in greater detail. Steward's work is emphasized in this historical section because, I will argue, most proponents of predictive locational modeling adopt—though not always consciously—both a cultural ecological position on the nature of culture and the cultural ecological causal approach to understanding.

In the second major division of this chapter the development of archaeological settlement pattern studies is discussed as it relates to these developments in theory; many settlement pattern studies differ from predictive locational models only in their lack of explicit extrapolation to a spatial population. This specialized discussion does not attempt to summarize the entire history of settlement pattern studies; see Parsons (1972) or Ammerman (1981) for a more comprehensive review.

Finally, the potential uses of predictive locational models from both management and research perspectives are set forth, followed by a few examples from the literature. These examples are meant to illustrate the diversity of approaches currently in use and some of the most obvious issues that these approaches raise. The reader interested in additional examples of recent locational models is referred to Kohler and Parker (1986) and to the appendix of this volume.

An important premise of this chapter is that predictive modeling as it is presently practiced is fundamentally about environmental determinism. That is why, in the next section, we briefly recapitulate the increasingly sophisticated forms this paradigm has taken. Why are the social, political, and even cognitive/religious factors that virtually all archaeologists recognize as factors affecting site location and function usually ignored in predictive modeling?

One obvious reason is that most models are constructed inferentially, starting from a sample of archaeological sites in a region and generalizing to an unknown

population of sites in that same region. This is made possible by resorting to maps displaying environmental categories across the total region with which site locations have been empirically correlated in the sample. At the same time, a total mapping of sites (the remains of the social and political network) is not available, or a predictive model would not be necessary.

Altschul is clearly correct when he says, in the next chapter of this volume, that "magnet sites" may significantly affect settlement density in their neighborhoods, presumably for reasons that go far beyond factors of the physical and biotic environments. In his example, the density of settlements around major Hohokam sites in the Santa Cruz River Valley of southern Arizona was greater than predicted on the basis of environmental features. And yet, it is possible to find examples in the archaeological record where precisely the opposite effect has been documented. In some periods of its history, for example, Teotihuacán in the Basin of Mexico seems to create a vacuum around itself; in others, sites seem to be denser in its vicinity than elsewhere (maps associated with Sanders et al. 1979). To further complicate matters, such changes may be due in part to changes in the area's role in a much larger, supra-regional system (see Paynter 1982:xi) that may be poorly understood. On a smaller, simpler scale, the large Pueblo I site of Grass Mesa in the Dolores River Valley of southwestern Colorado also seems to have created a partial settlement vacuum in its vicinity during the peak of its occupation (Kohler 1986:37).

This brings us to a second reason why nonenvironmental variables have not been used in most predictive locational models: archaeologists simply don't know *how* to use them. It is reasonable to believe that our sister disciplines, such as geography, might have solved such problems, particularly for the non-hunter-gatherer societies that they have emphasized. This is not the place for an exhaustive review of geography, but it is worth mentioning two approaches commonly used in the geographic literature to see whether they might help us.

One such approach with deep roots is the well-known central place theory, conceived by Von Thünen in 1826, expanded by Christaller in 1933, and introduced to the English-speaking world by Ullman in his famous article, "A Theory of Location for Cities" (1941; in Boyce 1980). Among other things, the theory predicts that cities will arise in the centers of productive areas; that they will be larger as their tributary areas become larger; that when a region is packed with cities the "tributary" spaces will be best described as hexagons; and most important, that a hierarchy of city size occurs, with centers in each class being predictable in number and in distance from each other. Ullman noted, as have many others, that the actual location of centers may be distorted by the distributions of resources and transportation routes, and that

> the type of scheme prevailing in various regions is susceptible to many influences. Productivity of the soil, type of agriculture and intensity of cultivation, topography, governmental organization, are all obvious modifiers. . . .
>
> The system of central places is not static or fixed; rather it is subject to change and development with changing conditions. . . .

> Christaller may be guilty of claiming too great an application of his scheme. His criteria for determining typical-size settlements and their normal number apparently do not fit actual frequency counts of settlements in many almost uniform regions [in Boyce 1980:176–177].

Given the subtleties and especially the fluidity of the sociopolitical environment, is it any wonder that archaeologists have chosen to concentrate on those relatively stable, "distorting" factors of the natural environment for locational prediction? In general, the central place model appears to be more valuable for *analysis* of a total spatial pattern of contemporaneous settlements than it is for *prediction* of the total distribution from some small subset of it. Nevertheless, it does have potential within a predictive context if enough of the settlement system is known to enable discernment of levels of size-class hierarchy, typical spacing of settlements within levels, and degree of influence of the various environmental factors serving to distort the ideal pattern. (For more discussion of central place modeling see Haggett et al. 1977 and various articles in Smith 1976.)

Another possibly relevant line of inquiry in geography is the study of industrial location. Let us assume for a moment that there are enough similarities between the problem of minimizing transport costs in the placement of factories and in the placement of relatively stable residential locations to make such an analogy worthwhile. Economic assumptions have thoroughly permeated this field so that, at least until very recently, profit maximization, in the context of perfect and complete information and thorough predictability of future circumstances, has been the single goal guiding analysis. In Alfred Weber's "least-cost" model (1929), the goal was to minimize transportation costs per unit of production, although benefits of agglomeration and labor availability might slightly distort the location predicted to be ideal on this basis (Gold 1980:217–231).

Later refinements of Weber's approach concentrated on correcting overly simplistic assumptions about transport costs, market demand, and methodological factors, and it was not until the 1970s that analysts began to question its reliance on economic factors in general and distance costs (as opposed to other costs) in particular (Gold 1980:218). Now, to judge by Gold's recent review of this area, interest centers on questions that were previously ignored, including how decisions are made in industrial organizations; how the wider industrial, business, and sociopolitical environments affect locational decisions; the extent to which attitudes about regions predispose locational behavior; and how locational searches are actually conducted. Gold concludes that such research is at an "exploratory stage" but that previous (exclusively economic) theory put forward "a model of behavior which, by its inherent assumptions, says little about the processes by which real-world locational decisions are reached" (Gold 1980:230–231). At this point it appears that archaeologists can profit from reading this literature but will not be able to find here a working, realistic model that will solve their own problems.

While archaeologists must redouble their efforts to build workable models with predictive power that take into account how social and political variables as well as those of the narrow economic environment affect location, and while

questions as to how locational decisions are reached in relatively small scale societies need more attention from archaeologists, the field as it presently exists—not as it perhaps should be—is the subject of the remainder of this chapter.

THE WIDER HISTORICAL DEBATE: HOW AND TO WHAT EXTENT DOES THE NATURAL ENVIRONMENT INFLUENCE HUMAN BEHAVIOR?

Attempts to explain differences among human societies are as old as the recognition of that diversity. The role of environmental factors in creating this diversity has been a subject of inquiry and debate since antiquity. In classical times these inquiries ranged from abstract questions about the origin of the earth and of humankind to the search for

> rational explanations for the existence of both health and disease, explanations which called for consideration, among other factors, of the nature and direction of winds, the effects of swamps and damp places, the relation of sunlight and of the sun's position in the heavens to the proper siting of houses and villages, and which, by extension, encompassed investigation of the effects of "airs, waters, and places" on national character [Glacken 1967:7-8].

An early example of this perspective is the *Histories* of Herodotus. Written in the fifth century BC and primarily concerned with the struggle of the Greeks to free themselves from Persian influence, the *Histories* also provides sketches of some 50 societies with attention to their geographic location, environment, dress, food, dwellings, form of self defense, and prestige as judges among other peoples (Hodgen 1964:23).

Yet in the Mediterranean world following the collapse of the Roman Empire, this comparative, cross-cultural tradition of inquiry that included environmental factors within its scope lost ground to theological interpretations of cultural diversity. Diffusion of the original Adamic culture, as outlined in the first chapters of *Genesis*, followed by local degeneration was generally considered to be sufficient explanation for diversity through the fifteenth and sixteenth centuries (Hodgen 1964:254-294).

A prominent dissenter was Jean Bodin, a French jurist writing towards the end of the sixteenth century, who argued that

> a sound solution to the problem of cultural diversification was not to be clouded by controversy over the early peopling of the world, or by a theory of original sin, migration, or the breakdown of tradition among the bearers of the Adamic tradition. Leaving all of this to one side, he elected to take man as a given, concentrating on the relation of several cultures to land, to climate, and to the topographical features of the several geographical regions. . . .

> The physical constitution of men, or their humoral makeup, determined their moral aptitudes or dispositions. Environment, climate, the conditions of time and place, did all the rest, reacting on men through their bodies [Hodgen 1964:276, 278].

For example, Bodin characterized people from hot climates in the northern hemisphere as being small of stature, weak, dark-haired and dark-skinned, fearful of heat, sad, hardy, mutinous, solitary, sober, and philosophic. People from cold regions were supposed to exhibit the opposite qualities (Hodgen 1964:279–280).

Grand schemes seeking to establish causal connections on ethnic, regional, or even continental scales between environmental factors (especially climate) and a wide variety of racial and cultural characteristics became more prominent in eighteenth-century Enlightenment thinking. Even the Baron de Montesquieu, although he was particularly prone to considering the form of government as the factor affecting all other aspects of society, did not ignore the influences of climate and environment. He was also willing to accord different factors causal primacy among different societies:

> Nature and climate rule almost alone among the savages [people with no nonlocal political structures and no domesticated plants or animals]; customs govern the Chinese; the laws tyrannize in Japan; morals had formerly all their influence in Sparta; and the ancient simplicity of manners once prevailed at Rome [Evans-Pritchard 1981:7].

We may conclude that even in the humanistic, rationalistic eighteenth century some natural philosophers took the position that, at least for some societies, causal initiative was to be found in the natural environment rather than in the mind. It was in reaction to such views that towards the end of the eighteenth century John Adams was led to complain,

> The world has been too long abused with notions that climate and soil decide the characters and political institutions of nations. The laws of Solon and the despotism of Mahomet have, at different times, prevailed at Athens; consuls, emperors, and pontiffs have ruled at Rome. Can there be desired a stronger proof, that policy and education are able to triumph over every disadvantage of climate? [Glacken 1967:685].

Montesquieu in particular, and to a lesser extent some of his contemporaries, clearly saw the interrelationship and interdependency among all aspects of a society (Evans-Pritchard 1981:4), thus laying the foundations for a functional view of culture that is one of the building blocks for modern cultural ecology. Although sweeping generalizations establishing connections directly from climate to human personality sound remarkably odd today, they represent unsophisticated precursors to modern cultural ecological positions that differ mainly by invoking a more credible and restrained chain of causation.

On the whole, however, eighteenth-century environmental or geographic determinism was a minor thread in a fabric that

> stressed the factor of conscious rational choice as the key to explanation of sociocultural differences. . . . [Enlightenment theoreticians] could not see a superorganic system interacting with the natural environment and responding with adaptive evolutionary transformations, which were neither comprehended nor consciously selected by the individual members of the society [Harris 1968:51].

A necessary prerequisite to the techno-environmental perspective as espoused by Harris was a credible theory of evolution, supplied for biology in the mid-nineteenth century by Darwin, even as Spencer was elaborating a similar theory for sociocultural evolution—a theory already expressed in part in the earlier writings of Turgot, D'Holbach, and others (Harris 1968:123). The goal of the great anthropologists over the last half of the nineteenth century (Spencer, Tylor, and Morgan) was to develop cultural evolutionary sequences using data from archaeology and from contemporary primitive societies. Their comparative method used modern "survivals" of earlier forms, not necessarily as exact replicas of stages through which other groups had progressed but as models from which something could be learned about earlier adaptations.

In geography at this time the focus of interest continued to be on the sort of geographic determinism espoused by Jean Bodin and reflected in some of the writings of Montesquieu. This view is strongly expressed in the writing of the nineteenth-century German geographer Friedrich Ratzel (1896–1898). Ellen Semple, who helped to interpret the ideas of Ratzel to the English-speaking world in the early 1900s and who is often regarded as an extreme geographical determinist, wrote of the effects of environment and climate on human stature, musculature, pigmentation, vocabulary, economy, population density, and migration, as well as of the "physical effects of geographic environment" (Semple 1911:40).

Interesting counterpoints to such views also appeared in the nineteenth century, however. The reciprocal nature of the relationship between people and their environments—ignored in simple environmental or geographic determinism—was beginning to be appreciated in some quarters. George Marsh, in *Man and Nature, or Physical Geography as Modified by Human Action* (1864), reasoned that many important influences emanate not from nature to humans but rather in the opposite direction.

In the early years of the twentieth century, historical particularism, most purely exemplified by Franz Boas, constituted a rebellion against the largely unilinear cultural evolutionary sequences of the nineteenth century and against the comparative method used by Morgan, Spencer, and others. Nor did this new school of anthropological thought have any use for the simple, mechanical, large-scale correlations among environmental features, race, and culture that were still being promulgated by some geographers. One of Boas's most prominent students evaluated the causes behind the historical particularists' avoidance of environmental factors in the discussion of cultural phenomena:

> In part this represents a healthy reaction against the old naive view that culture could be "explained" or derived from the environment. For the rest, it is the result of a sharpening of specific anthropological method and the consequent clearer perception of culture forms, patterns, and processes as such: the recognition of the importance of diffusion, for instance, and the nature of the association of culture elements in "complexes." Most attention came to be paid, accordingly, to those parts of culture which readily show self-sufficient forms: ceremonial, social organization, art, mythology; somewhat less to technology and material culture; still less to economics and politics, and problems of subsistence. Much of the anthropology practiced in this country in the present century has been virtually a sociology of native American culture; strictly historic and geographic interests have receded into the background, except where archaeological preoccupation kept them alive [Kroeber 1939:3].

Ironically, in his ethnographies Boas remarked on environmental factors influencing site location, as in his astute observation that the distribution of population among the Central Eskimo was strongly related to conditions of sea-ice favorable to hunting the ringed seal (Damas 1969:1). In his later, more general work, however, he downplayed the role of the environment as a determinant of human behavior.

Another ironic feature of the impact of historical particularism on anthropology is that it showed the way for a more productive analysis of the relationship between culture and environment. By reducing the scale of his observation—by being a particularist—Boas in some ways anticipated a more modern approach to the problem of correlating settlement practices with environmental features. In his discussion of the roots of ecological explanation in anthropology, Ellen (1982:5–6) makes the important point that

> The problem of drawing correlations between environmental and social phenomena is very much a question of magnitude—the geographic (or demographic) scale of the correlations postulated. . . . The more specific the correlation the greater the possibility of there being a single determining relationship and the greater the accuracy in predicting future events under specified conditions.

This is a crucial observation for the task of locational modeling. Many valid criticisms can be made of naive environmental determinism for its suggestions of large-scale, simplistic correlations between environmental and cultural features. These criticisms are not all germane, however, to more specific correlations between *certain* environmental features and *certain* aspects of human behavior. Settlement systems and ecosystems are both complex, and we should not expect to find simple correlations between them. The task of locational modeling is to isolate those aspects of the environment that do influence settlement behavior and place them into perspective with nonenvironmental factors that also influence settlement behavior.

In the generation of anthropologists following the period in which historical particularism reached its ascendancy, people like Kroeber and, to a lesser extent, Wissler (e.g., Wissler 1922) once again began to study the relationship between environment and culture. This time, however, the relationship was stripped of causality. Both Kroeber and Wissler were interested in culture areas that were

> relativistically defined in terms of their distinguishing characteristics and occurred in different environmental settings. . . . [T]he concept of adaptation of the cultures, especially of the nature of social groups . . ., [was not] taken into account. In fact, this would smack of reductionism, which Kroeber, holding firmly to the idea that cultures should be dealt with on the superorganic level alone, had always opposed [Steward 1973:53–54].

Julian Steward and Cultural Ecology

The contributions of one of Kroeber's students ultimately have had more impact on archaeology than those of Kroeber himself. Along with a number of influential contemporaries that included Omer Stewart and Leslie White (Stewart 1943; White 1949), Julian Steward (1938, 1955) was responsible for three advances in the discussion of environmental concepts that have specific importance for the practice of locational modeling. First, Steward, unlike anthropologists using the culture-area concept, was interested in causal explanation rather than correlation; second, he emphasized the effect of particular *local* aspects of the environment on particular facets of culture, thus moving away from large-scale correlations of regional environments with "culture types"; third, he identified more or less specific pathways through which environments might influence cultures (in his "culture core" concept) and tried to devise a procedure for studying the extent of these influences (Ellen 1982:52–53).

In Steward's terms, those aspects of a culture that were most closely connected with environmental exploitation constituted the "culture core"; other aspects, determined by purely cultural historical factors, were considered secondary features. Core features and secondary features had to be identified empirically, and these could be expected to differ in differing environments and cultures. For a particular culture, discrimination between core and secondary features began with an examination of the natural environment and of the relations between the environment and the economy. Next, the patterns of behavior involved in exploiting this environment with a specific technology were recognized. Finally, the influence of these behavior patterns on other aspects of culture was assessed (Steward 1938:2; 1955:37, 40). All aspects of culture implicated in these investigations constituted the core; the residua were the secondary features. This procedure clearly reveals the direction and type of causality that Steward believed to be at work in the relationship between environment and culture.

Not all features of the natural environment equally influence the core of culture, and what *is* important may be expected to vary from area to area. For the aboriginal groups occupying the Great Basin and adjacent portions of the Columbian and Colorado plateaus at the time of contact with Euroamericans, for example, Steward suggested that "the important features of the natural environment were topography, climate, distribution and nature of plant and animal species, and, as the area is very arid, occurrence of water" (1938:2). He took the density and distribution of the population; the division of labor at sexual, familial, and communal levels in hunting, fishing, and seed-gathering; the territory covered and the time required

for different economic pursuits; and the size, composition, distribution, and degree of permanency of villages to be behavior patterns that were directly and strongly influenced by the nature of the environment, in the context of the technology available to exploit it (Steward 1938:2).

His comments on the village locations of specific groups were based on conversations with informants who were recalling a lifeway that by that time was extinct and, usually, on visits to the areas in question. Many of these comments indicate which factors Steward considered to be determinants of site location. The Northern Paiute of Owens Valley, for example, lived in an area that was rich and diverse in comparison with most of the Great Basin. Their villages were relatively permanent and were situated on the alluvial fans of streams where these watercourses emerged from the canyon wall, about 2–4 mi from the Owens River. These locations afforded access to abundant water and were centrally located with respect to critical floral resources (except for piñon nuts) growing in or near the valley. Sites related to piñon nut extraction and use were located in the adjacent Inyo and White mountains and might be occupied during part of the winter in the event of abnormally abundant harvests. As important determinants for winter (or permanent village locations for all the groups he studied, Steward repeatedly mentions the availability of water, ample timber for houses, and fuel, and he also emphasizes avoidance of areas with unacceptably cold winter temperatures. Thomas (1973) used simulation to predict what the artifact dispersal patterns should be if Steward's reconstruction of the Great Basin Shoshonean subsistence-settlement system applied to precontact times in the Reese River Valley. Steward's predictions, as operationalized by the simulation, were generally verified.

Some of Steward's views on the responsiveness of site location to environmental factors will be systematized into a more general framework in Chapter 4 and hence are worth additional discussion here. To judge by Steward's work, the locations of winter villages in the Great Basin ought to be relatively predictable on the basis of associated environmental features. For example, Steward characterized the entire Shoshonean culture as practically, even "gastrically," oriented. Since Shoshonean groups were frequently at risk of starvation, their adaptation (broadly speaking, including the location of their settlements) was constantly exposed to selective processes. Social and political factors that may affect site location—defensibility; access to trade partners and routes; and economic, social, and political obligations to nonlocal groups—were of minimal importance in comparison with many areas in North America where warfare was more frequent, economic specialization more pronounced, the family not the basic economic unit, and social and political groups more rigidly structured and less local. It will be argued in Chapter 4 that this constellation of factors—which will be placed on the low end of a continuum of "intensification"—results in settlement behavior that is quite responsive to environmental factors. Moreover, the structure of the environment is such that the resources apparently affecting winter village location are relatively concentrated in space, overlap to a fairly high degree, and exhibit either high temporal constancy—meaning that they can always or nearly always be found in the

same place, as in the case of water and certain aquatic resources—or high temporal contingency—meaning that they are seasonally predictable. It will be argued in Chapter 4 that this kind of patchiness and this kind of temporal predictability make for high site visibility and high site predictability on the basis of environmental variables.

The location of piñon-gathering stations, on the other hand, depends in part on the distribution of piñon resources, which in any year are relatively widely distributed, seldom overlap with other critical resources, and exhibit low temporal predictability. Logically, this environmental structure should lead to dispersed, poorly visible, and poorly predictable distributions for archaeological materials deposited during piñon exploitation. On the basis of these observations, and of Steward's discussions, we would expect different parts of this settlement system to have differing visibility and variable degrees of predictability on the basis of environmental variables.

Even this brief discussion of Steward's approach and conclusions clarifies the continuity between inductive locational modeling and Steward's work. Steward demonstrated that—at least for some site types and in some environments exploited by some groups in the arid portions of western North America—there is good reason to believe that location was highly responsive to a relatively limited number of map-readable environmental determinants. In addition, he argued for a more or less one-way directionality of influence: from the environment, as exploited by a particular technology, to the culture core. Finally, although his research was influenced by a strong and consistent theoretical orientation, Steward argued that the particular aspects of the environment that are most relevant to adaptation (which is to say, to the composition of the culture core) have to be discovered empirically.

People in Their Ecosystem: Post-Stewardian Developments

Locational modeling—particularly in its inductive variety—normally assumes that certain environmental variables strongly influence site location. If settlement behavior can be considered to be part of the "culture core," this assumption finds support in Steward's cultural ecology. The strong, although frequently implicit, reliance of locational modeling on Steward's theories or on other variants of what Trigger (1971) calls "deterministic ecology" makes the resultant models susceptible to the many criticisms to which Steward's work has been subjected in the last two decades.

One outstanding problem is an ambiguity in the definition of the culture core, which is noted both by Harris (1968:660–662) and by Kohl (1981:102). There is no rigorous objective procedure for determining what constitutes the core, and it is clear from Steward's own statements that the core may occasionally encompass social, political, and even religious patterns. May we assume that all aspects of settlement behavior are core elements? If not, which aspects are? Another problem

is Steward's assumption of an unrealistically unidirectional influence of the environment on culture. A third problem is unrecognized complexity and variability in how the environment is perceived in different cultures (Brookfield 1969).

Steward's approach enjoys continued popularity among many practicing archaeologists, especially those involved in hunter-gatherer studies (Bettinger 1980:190). As a result of these problems, however, and perhaps also as a result of the increasingly sophisticated ecological studies of the last two decades, many human ecologists and some archaeologists have begun to abandon Steward's framework in favor of an *ecosystemic* perspective influenced by evolutionary ecology—a development that is more evolutionary than revolutionary. A very selective sample might include publications by Marston Bates (1953), J. W. Bennett (1946), Harold Brookfield (1968), J. G. D. Clark (1952), David L. Clarke (1968), Harold Conklin (1961), Kent Flannery (1968), Stanton Green (1980), Donald Hardesty (1975), Robert Netting (1974), Roy Rappaport (1971), and Bruce Winterhalder (1981), among many others. Although each of the researchers who has shifted to an ecosystems perspective has unique points to make, Roy Ellen (1982:75–78) has attempted to summarize several characteristics shared by most workers involved in this reorientation of culture/environment studies:

1. *Monism.* Behavioral and environmental traits are analyzed as part of a single system. Culture becomes part of animal behavior, or at least it must follow rules that do not contradict those imposed by natural selection.

2. *Complexity.* Significance and causality in this single, integrated system containing both the culture and the environment are "found in the web of finely interrelated factors rather than with general propositions at the level of gross categories" (Ellen 1982:76).

3. *Connectivity and mutual causality.* "In the ecosystem view, all social activities impinge directly or indirectly on ecological processes and are themselves affected by those same processes. Fauna (including humans), vegetation, soil structure, and microclimate are intricately related and mutually interdependent (Ellen 1982:76).

4. *Process.* In this systemic view of relationships the emphasis is on the interaction of variables (for example, positive and negative feedback relationships) rather than on correlations between social and environmental variables at particular states of the system.

5. *Populations as analytic units.* Local human populations replace societies as units of observation and analysis, a situation analogous with the ecological analysis of nonhuman populations.

Local, detailed paleoenvironmental reconstructions are of special concern to the archaeologists involved in this reorientation, and this is a concern with which Steward would have been sympathetic. There is an increasing awareness that such information must not simply be brought in as an after-the-fact explanation for observed changes through time in human use of the landscape, as has long been the

practice. Rather, settlement system studies should account in a dynamic manner for changing resource distributions related to changing climates (e.g., Darsie 1983).

The challenge to Steward's approach posed by these advances is also implicitly a challenge to locational modeling as typically practiced. Future advances in locational modeling depend on our learning how to incorporate the richness and complexity of the systemic perspective in our locational predictions.

THE EMERGENCE OF SETTLEMENT PATTERN STUDIES IN ARCHAEOLOGY

One important result of Julian Steward's insistence on the importance of the local environment in the study of living (and recently living) cultures and of his interest in the location of ethnographic settlements was the development of studies of archaeological settlement patterns. The survey component of the Virú Valley program conducted in the late 1940s was instituted largely as a result of Steward's influence (Willey 1953:xviii). Willey's 1953 monograph about this work is generally regarded as having defined a new field of inquiry in archaeology:

> The material remains of past civilizations are like shells beached by the retreating sea. The functioning organisms and the milieu in which they lived have vanished, leaving the dead and empty forms behind. An understanding of structure and function of ancient societies must be based upon these static models which bear only the imprint of life. Of all those aspects of man's prehistory which are available to the archaeologist, perhaps the most profitable for such an understanding are settlement patterns.
>
> The term "settlement pattern" is defined here as the way in which man disposed himself over the landscape in which he lived [Willey 1953:1].

Willey included within the scope of settlement pattern studies the nature of dwellings and their arrangement within settlements and the nature and distribution of communal buildings. His discussion of the role of environmental, technological, and demographic change in affecting settlement patterns is not elaborate by modern standards; he was much more interested in how the community patterns of these large, late-prehistoric sites in Peru were affected "by various institutions of social interaction and control" (1953:1).

Nevertheless, a field *was* defined, and a series of papers (Willey 1956) published three years after Willey's Virú Valley report contains many contributions emphasizing the importance of environmental variables in determining the distribution of human populations across the landscape (e.g., Haury 1956; Heizer and Baumhoff 1956; Williams 1956). Other authors (e.g., Sears 1956) were interested more in the social and political aspects of community patterning than in environmental relations. In 1968 Trigger defined the various aspects of settlement patterns somewhat more rigorously than had previously been done, and he distinguished among the probable determinants of location for individual buildings, community layouts, and

"zonal patterns" (Trigger 1968). In the case of zonal patterns, he states that "the overall density and distribution of population of a region [are] determined to a large degree by the nature and availability of the natural resources that are being exploited" (1968:66). He notes, however, that broad economic (as opposed to simple subsistence), political, religious, and defensive factors may also be important determinants of site location among agriculturalists.

In the 1970s several important initiatives added new items to the list of environmental variables that archaeologists were willing to consider as possible determinants of location, and they also affected the ways that these variables were handled analytically. For example, the "situation" of a site (Roper 1979:11–14) or the putative "territory" of the community occupying it (Vita-Finzi and Higgs 1970) began to be scrutinized in addition to the more traditional on-site environmental characteristics. Catchment analysis, as this investigation is usually called, was designed to provide insight into the economic activities of the occupants of a site. Like most efforts to use the distribution of environmental variables in understanding site location, catchment analysis makes the joint assumptions that

> the most important transactions for most people were with the environment . . . [and that] humans tend to minimize the time or effort expended in their economic transactions with the environment (or perhaps they *include* effort and time expenditure as considerations in these transactions). In societies without advanced transportation these two factors—strong economic coupling with the environment and minimization of time and effort—encourage location close to important economic resources [Kohler and Parker 1986:400; emphasis original].

Another important advance made in the 1970s was in the analysis of data. Steward himself had avoided statistical approaches, and following perhaps unconsciously in his footsteps virtually all settlement pattern studies for many years followed an anecdotal form. That is, the investigator called attention to apparent tendencies for sites to be located in areas having specific constellations of natural features, much in the same way that Steward did in his Basin-Plateau work cited above. Where these relationships were patent, the observations were probably correct, at least to the extent that the original surveys were not biased by an internalized model of "where sites should be." Nevertheless, it was a great contribution to settlement pattern studies when the participants in the Southwestern Anthropological Research Group (SARG) helped to introduce a more rigorous testing procedure for determining the degree of relationship between site locations and environmental variables. This procedure involves the creation of expected site distributions for comparison with observed site distributions, using formal statistical inferential techniques.

The SARG organization was dedicated to investigating systematically the question of why archaeological sites (or, in some versions, prehistoric population aggregates) in the Southwest were located where they were (Plog and Hill 1971). The members of SARG began with the basic assumption that activities were located in such a way as to optimize the return on energy investment and then proposed

three somewhat more specific hypotheses for testing. These hypotheses suggested that activity loci were

1. situated with respect to critical on-site resources,
2. situated so as to minimize the effort expended in acquiring required quantities of critical resources, and
3. located so as to minimize the cost of resources and information flow among loci utilized by interacting populations (Plog and Hill 1971:12).

Most participants concentrated on the first two problems, and in his perceptive insider's view of the SARG research design several years after its inception, Dean (1978:107) suggests that this was due to procedural and logistical considerations. The difficulty of operationalizing and testing the third hypothesis would have been great.

Plog and Hill's suggested procedures for testing these hypotheses using null models and statistical comparisons of where sites were and were not located were rarely used by the SARG participants. More often, the SARG researchers concentrated on searching for significant differences in site location frequencies across environmentally defined strata. The methods proposed by Plog and Hill have, however, become standard in cultural resource management and in some research contexts. The potential utility of this brand of locational research was clearly foreseen by Plog and Hill (1971:11):

> our research should lead to the ability to predict site locations (and something about organizational characteristics of sites) from the distribution of critical resources and other critical variables. And, conversely, we ought to be able to predict the critical variables by examining the site distribution patterns.

Some of the problems with the "critical resources" concept are noted in the Chapter 4 discussion of how variables are selected—in inferential or deductive models—as potential determinants of locational behavior. Hill (1971:58) suggests that critical resources are those "without which the system would collapse" (but see Sullivan and Schiffer 1978:172). Dean (1978:108) acknowledges that SARG has been primarily concerned with food resources and suggests that availability of fuel, structural wood, and other nonfood resources might also be important in determining site location.

While it is clear that those of us who are engaged in locational modeling owe a substantial debt to the SARG participants, it is important to call attention to a final comment by Sullivan and Schiffer concerning the difference between investigating the distribution and movement of people through space in the systemic, behavioral context and investigating the spatial distribution of archaeological sites:

> [P]rehistoric peoples most likely did not locate "sites" anywhere. However, they did establish, occupy, and abandon behaviorally significant spaces, such as activity areas, camps, and settlements. . . . Sites are nothing but deposits of material remains in the environment that archaeologists recognize as being potentially informative about past

> cultural behavior and organization. . . . Owing to secondary deposition, multiple occupations, and other formation processes, sites usually are not equivalent on a one-to-one basis to camps, settlements, or population aggregates [Sullivan and Schiffer 1978:169].

The discovery of statistical associations between site types and environmental variables, they continue, may be potentially useful for developing predictive models for cultural resource management (CRM) and for evaluating survey samples, but construction of such models "has little to do with the formulation and testing of behavioral principles" (1978:169).

THE ERA OF PREDICTIVE MODELING

It is clear from the above citations that in the early 1970s there was already some talk about predictive modeling, although there were relatively few examples of what this term might mean. To avoid ambiguity, we can define a predictive locational model as a simplified set of testable hypotheses, based either on behavioral assumptions or on empirical correlations, which at a minimum attempts to predict the loci of past human activities resulting in the deposition of artifacts or alteration of the landscape. Thus defined, the potential applications of predictive models are certainly not limited to CRM contexts. Green (1973) conducted a locational analysis of prehistoric Mayan sites (defined as the loci of one or more structures) in northern British Honduras (now Belize). In this research she shared the SARG assumption that "sites were located so as to minimize the effort expended in acquiring critical resources" (1973:279). Several soil and vegetation variables, along with variables reflecting distance from navigable bodies of water (in the belief that access to commerce was a critical resource), were tested for association with counts of sites per unit area, using multiple linear regression. The resultant multivariate statistical model of site location was interpreted as predicting high probability for site location in areas with large tracts of good agricultural land and in proximity to trade routes. In a sample of 150 quadrats known to contain only 22 sites, about 22 percent of the variance in the number of sites observed in each 4.25 km^2 quadrat was explained by the independent variables selected by the regression routine. Quadrats with high negative residuals (no sites found, several predicted) were considered as probably containing undiscovered sites, and such quadrats were assigned a high priority for future survey efforts. Because sites were located in the centers of arable tracts rather than on their margins, Green inferred that residences were probably located so as to have garden plots in their immediate vicinity.

As predictive models began to be applied in CRM contexts, many still-unresolved issues concerning the appropriate use of predictive models were identified almost immediately.

> Predictive models are probability statements; they are not "facts," and cannot substitute for facts in any application requiring the use of hard data about specific individuals as decisionmaking criteria. . . .
>
> The problem is that some archaeologists have told some planners that our predictive models can be used as hard data, when in actuality it is our hard data on site location and significance that must be figured into the planner's cost-benefit ratio. To substitute a scientific hypothesis (our predictive model) for scientific fact (actual site location) as a criterion for a planning decision is to court disaster.
>
> There is only one way for us to get the hard data for use in such decisions: by an intensive ground reconnaissance of the entire area to be affected by a proposed project [Wildesen 1974:1-2].

In the latter half of the 1970s the Bureau of Land Management, Forest Service, Corps of Engineers, Interagency Archeological Services, and some State Historic Preservation Officers were beginning to sponsor both surveys that would result in predictive models and attempts to build predictive models from data already collected (Interagency Archeological Services [IAS] 1976:3; King 1978:73). Although important federal historic preservation legislation dates back to the turn of the century (the Antiquities Act of 1906; the Historic Sites Act of 1935), the National Historic Preservation Act of 1966, amended in 1976 and 1980, has been of signal importance in this growth of predictive models, especially Section 106 of that act, which requires that federal agencies "take into account" the effects of their actions on properties eligible for the National Register of Historic Places (King 1984; Scovill 1974). In conjunction with Executive Order 11593 (1971), other sections of the National Historic Preservation Act, the National Environmental Policy Act of 1969, and various implementing regulations, this statute gives federal agencies the "substantive responsibility to identify historic properties on their lands and nominate them to the National Register, and to record such properties when they must be destroyed" (King 1984:116). Highly variable legislation for the protection and identification of archaeological resources also exists in state and local jurisdictions (Rosenberg 1984).

Federal (and occasionally state) agency response to this legislation has included predictive modeling, under the assumption that it will be a long time (to say the least) before a total, comprehensive inventory of archaeological resources can be conducted on lands under their jurisdiction.

> For comprehensive planning, predictive survey may best be considered an ongoing process in which increasingly fine-tuned predictions can be made as more and better information becomes available. If the archaeologist continues to survey a new selection of sample units every time, he will eventually obtain a 100 percent sample. This is a rational goal for statewide comprehensive surveys and for federal agency surveys conducted under section 2(a) of Executive Order 11593. The advantage of predictive survey is that *some* useful data for purposes of planning in the entire study area became available almost immediately . . . and it is probable that all the information needed to carry out responsible preservation planning will be available before physical inspection has covered even 50 percent of the land [King 1978:92; emphasis original].

The flood of predictive models that appeared in the late 1970s shows that contractors were happy to respond to agency requests for such models, even though (judging by the variability in techniques and products) no one was sure how prediction might best be accomplished. Early attempts include Dincauze and Meyer (1976), Fuller et al. (1976), Hackenberger (1978), Robertson and Robertson (1978), Scott et al. (1978), Woodward-Clyde Consultants (1978), Holmer (1979), Barber and Roberts (1979), Burgess et al. (1980), Kohler et al. (1980), Muto and Gunn (1980), and Senour (1980).

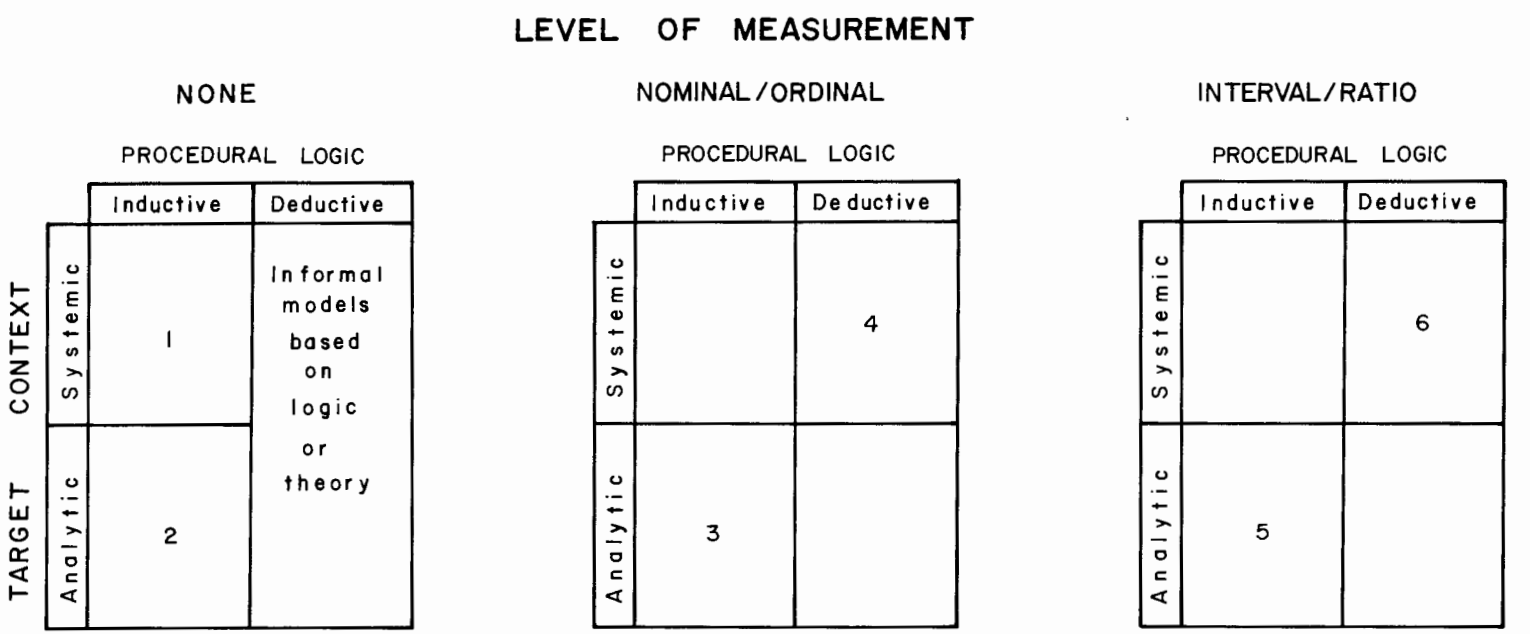
A Taxonomy for Predictive Locational Models

Before we can begin to talk about the very dissimilar enterprises that have been called "predictive locational models" during the last 10 years, we need to establish some definitions and build a classification for what has been done so far. Another purpose for classification is to highlight what this author believes to be the most significant dimensions of variability among the predictive locational models put forward to date. Specifically, I propose a classification with three distinguishing dimensions: level of measurement, procedural logic, and target context (Figure 2.1).

Many models for site location or settlement behavior are intuitive or not fully operationalized. The ugly word *operationalization* refers to the process of careful definition of all the terms in a model in such a way that the same predictions can be made from a model by different people. If a model can be objectively, replicably mapped, it is operationalized; a model consisting of the statement that "sites are located near rivers on dry, level ground," for example, is not mappable until site, near, river, dry, level, and ground have been rigorously defined.

As we move to the right in Figure 2.1, we move from models with no measurement to models based on variables measured at the categorical or nominal level (such as soil type) or ordinal level (such as resources ranked in order of hypothesized importance) to models based on variables measured at the interval or ratio level (such as slope, distance to water, estimated net primary productivity, and so forth). There is nothing wrong with site location models that are not operationalized if they provide insights into settlement behavior, as does Binford's (1980) distinction, based on a review of hunter-gatherer subsistence and settlement system organization from around the world, between foragers and collectors. Until a model is operationalized, however, it cannot be mapped and cannot be used for management. This is one problem with the informal models of settlement pattern that are found in many Class I overviews based on existing literature and site files. The most important distinction along the dimension labeled *level of measurement* is between the box on the left, containing unoperationalized models, and the two boxes on the right, containing operationalized models.

The other two dimensions in this classification—procedural logic and target context—need to be discussed together. Most predictive models in cultural



Examples:

1. Many unquantified discussions of prehistoric settlement systems in particular regions; also, Binford's (1980) forager/collector model
2. Many unquantified discussions of prehistoric settlement patterns in Class I overviews
3. Pilgram 1982
4. Limp and Carr 1985
5. Most cultural resource management predictive locational models: e.g., Kvamme (1980), Nance et al. (1983)
6. Models based on optimal foraging theory (e.g., Winterhalder 1983:207-208) and other model-based approaches (e.g., Jochim 1976)

Figure 2.1. A suggested taxonomy for the different types of locational models that appear in the literature.

resource management have been *inductive* (used here synonymously with the terms *inferential* or *empirical/correlative*) in their logic. That is, they begin with survey data on the distribution of archaeological materials across the landscape in relation to some (usually environmental) features, and then they estimate the spatial distribution of the population of archaeological materials from which the sample was drawn. The logical alternative to this procedure is to begin with a theory as to how people use a landscape and to deduce from that theory where archaeological materials should be located.

By target context I mean the "theater of operations" for the model. The systemic context (Schiffer 1972) is the dynamic living system observed by ethnographers and ethnoarchaeologists. (Of course, it too is subject to inference, partial observation, and informant perception.) The sum total of the materials collected, altered, organized, and deposited by the participants in this system, and the spatial distributions of these materials, constitute the archaeological context (Schiffer 1972). This context can never be directly observed, however, and as soon as we begin to sample materials from it, analyze them, and make interpretations, we enter the interpretive or analytic context (Kohler et al. 1985). Some of the processes and activities in each of the contexts are discussed in Chapter 4.

In two senses inductive models automatically operate in the analytic context. First, to make predictions directly about the systemic context they would have to make some attempt to control for the effects of the postdepositional and depositional processes that separate the analytic from the systemic context (see Chapter 4); this is rarely, if ever, done. Second, and more insidious, the sampling and analysis processes of the analytic context are invisibly imbedded in inferential predictive locational models. Any inferential locational model predicts only what would have been found had the population of space from which the sample was drawn been surveyed in the same manner as was the sample, using the same rules for attribute coding, site recognition, and data analysis. Such inferential models predict neither the systemic interaction between a cultural system and a landscape nor the archaeological context resulting from it; rather, they predict what we will find and how we will interpret it if we consistently follow a particular set of rules for fieldwork and analysis. For this reason I say that inductive models normally operate in the analytic context. The challenge for inductive models is to build the bridge to the systemic context by making the analytic methods (including discovery) as "transparent" (non-bias-making) as possible and by controlling for the effects of depositional and postdepositional processes in the archaeological context.

Deductive models, on the other hand, begin with some theory predicting human behavior in the systemic context. The challenge for deductive models is to build the bridge to the analytic context, which is where the outputs of the system can be observed. This bridge-building—whether from the systemic to the analytic context or vice versa—is referred to as explanation (see discussion in Chapter 4). Explanatory models, as I suggest the term be used, are inherently neither inductive nor deductive. Instead, they are models that attempt to build the bridge between the dynamics of the living system and its observed outputs.

There is at least one sense, however, in which deductive models are clearly preferable to inductive models. Except when we are working with living groups we are limited to testing predictive locational models in archaeology in the analytic context. Thus, an explanatory, inferential locational model would end up making predictions about behavior in the systemic context that could not be immediately tested, although in a cycle of scientific inquiry these predictions could be used to suggest theory from which implications for future testing are drawn. An explanatory, deductive locational model would result in predictions for the analytic context that would be directly testable.

Examples

A detailed review of even a small proportion of the predictive models of the past decade would take much more space than is available here. The only reasonable way to approach this mass of material is to pick a few themes and trace them through a highly selective sample of the available references. Discussions of sampling, statistical methods, use of remote sensing data, and use of geographical information systems are generally avoided here, as they are treated in detail elsewhere in this volume. The four models to be discussed here were chosen because they illustrate particular cells in the proposed taxonomy and because they focus on various geographic regions.

I would suggest that some of the same criteria used to evaluate research designs and theory can be used to assess predictive locational models. One obvious criterion that should be applied is the accuracy of these models. Do they supply reliable predictions? Unfortunately, this information is available for so few models (see Appendix) that other, more general guidelines need to be considered. This, in itself, underlines the need for additional attention to model testing, refinement, and verification. In the discussion below of examples of predictive models, I have followed Blalock's (1979) suggested criteria for judging what constitutes good social science theory in general.

1. *Generalizability.* Generalizable models can be applied to large areas, rather than small; are applicable to different adaptations and environments, rather than just to one; take into account the entire settlement system, rather than just part of it; and have implications for human organizational systems in general as well as prediction of site locations in particular. Generalizability has both a conceptualization component—are the theoretical arguments applicable across a broad range of situations?—and a comparability component—if our theories can be applied across a broad range of situations, can our measurement operations be guaranteed to be applicable in the same broad range of circumstances (Blalock 1982:29)?

2. *Simplicity.* Other things being equal, a simple (or parsimonious) model is to be preferred to a complex one. After all, one reason people make models in

the first place is that the real world is too complex to be readily and unambiguously understood.

3. *Internal Consistency.* Like other models, predictive locational models must be mathematically and logically consistent.

4. *Precision.* Precision refers to the fineness of detail in the predictions. Precision may involve spatial detail: are predictions made to the square mile or the square meter? Or it may involve content: how fine-grained are the predictions of what will be found in various locations? Are various possible site types, periods, or assemblage types differentiated? Other things being equal, a model that is precise in its predictions is to be preferred to one that is not.

5. *Falsifiability.* It must be possible to prove that a model is wrong.

The last two of these characteristics can be lumped together for convenience, since a model that is not precise in its predictions cannot be falsified. Internal consistency is a more or less mechanical problem that needs no further mention here (but see Kohler and Parker 1986:398). There are, however, severe and perhaps unresolvable conflicts among generalizability, precision, and simplicity in predictive modeling, as in social science theory in general (Blalock 1982:27–31).

A Predictive Land-Use Model for North-Central Washington

In an overview based on a survey of existing literature, Robert Mierendorf et al. (1981) first constructed a predictive model for site location in a large study area encompassing the corridors of two proposed transmission lines and then carried out a "sensitivity analysis" for the predicted archaeological resources in these same areas. The sensitivity analysis was designed to predict the likelihood that disturbances in different geographic zones will significantly impair the research value of predicted archaeological resources, given the predicted regional research value of these resources, their density, and previous disturbances in each zone. I will consider only the predictive aboriginal land-use model in this discussion.

If we have to fit this model into one of the pigeonholes shown in Figure 2.1, it would probably be best to call this an inductive model aimed at the analytic (archaeological) context, at a nominal level of measurement, although to the (rather large) extent that the authors rely on an ethnographic model, it could be argued that this is primarily a deductive approach. There is no formal statistical model for site location, type, or density, but the model was operationalized to the extent that a map could be made. To the extent that the model construction relied on data from archaeological site excavation and survey, it is fair to call it inductively based. The model also takes into account the seasonal distribution and density of resources, however, and draws on recent hunter-gatherer studies. In some places it apparently (and implicitly) assumes a least-cost solution to location of settlements in cases of conflicts between the location of resources. For example, many researchers assume that storage of fish and roots was necessary in order for human inhabitants in the Columbian Plateau to survive the harsh, resource-poor winter months in a rela-

tively sedentary fashion. In this study area, fishing and presumably fish and perhaps root storage were concentrated along the large rivers, the Columbia and Okanogan. These same river valleys, however, were probably unfavorable winter locations from the point of view of adequate shelter from severe winter winds and the availability of wood for fuel. Mierendorf et al. assume that in decisions about winter village locations priority would be given to distributions of landforms providing shelter from winter winds and to the availability of fuel, which is a bulky, heavy item in comparison with stored food.

The predictive model is based on a vegetation map and a set of topographical contour maps. The model recognizes six broad zones of archaeological resource types and densities (Mierendorf et al. 1981:90):

1. Summer hunting and gathering zone; low density. Areas supporting summer hunting of dispersed ungulates. The highest elevations, which have a mesic vegetation and are accessible only in the summer, are mapped as part of this zone.

2. Summer and fall hunting and gathering zone; low density. Areas supporting dispersed ungulate hunting in the summer and fall. Intermediate elevations with a xeric vegetation are included in this zone.

3. Spring, summer, and fall (on map) and winter (in text) hunting and gathering zone; low density. Low-elevation, steppe vegetation zones not included in any of the other categories are mapped in this zone. These areas are relatively accessible in winter.

4. Summer fishing camp zone; high density. Areas within 10 km (6.2 mi) of falls and rapids on the Columbia and Okanogan rivers and mouths of tributaries to these rivers are included here. Catchment sizes are modified to reflect steep river valleys, resulting in a linear distribution for this zone.

5. Winter residence zone; moderate to high density. Areas in which stands of timber, protected canyons and valleys, and water resources are available within a 5 km (1.3 mi) radius of each other are mapped in this zone.

6. Overlap of fishing camp and winter residence zones; high density.

Generalizability. This model is intended to be applicable to a study area in north-central Washington that covers more than 21,000 km^2 (8000 mi^2). The temporal scope of the model is assumed to be the entire local prehistoric sequence. Its applicability to other areas may be slight, inasmuch as it relies on local ethnographic analogs and archaeological data for its predictions.

Simplicity. The model is moderately parsimonious in its selection of independent (causal) variables. Three different types of variables (shelter, fuel, and food resources) are considered for their possible effects on the locations of three different site types. Both on-site and catchment-area variables are considered. The model gains simplicity but loses realism and precision by not incorporating changing resource distributions due to changing climates and changing adaptation types due to intensification.

Precision. The model gains precision by considering seasonal distributions of resource types and by identifying differing site types and densities. On the other hand, the very large study area, the rather poor quality of available maps of important resource distributions, and the hand-measurement techniques all contribute to low spatial resolution in prediction. It is hard to imagine, particularly, how the distribution of the winter village zone could be accurately mapped using these manual techniques. The authors themselves call attention to these shortcomings (Mierendorf et al. 1981:84, 94).

In many ways this study is exemplary among the "overview" documents that attempt to predict prehistoric land use. Most such overviews result in unoperationalized models that remain at a verbal, unmapped, unmeasured level, somewhere in the far left-hand box in Figure 2.1. It also avoids too heavy a reliance on existing survey records that (if typical of most areas) are biased toward certain types of sites. This is achieved by giving more weight to natural resource distribution than to the existing site data base and by building a reasonable model for the use of those resources by using the ethnographic record. Even granting unlikely climatic stability assumptions resulting in unchanging resource distributions, the danger in such an approach, of course, is that if adaptation types other than those present in the documented ethnohistory were ever present, they will not be identified or predicted by such a model.

A weakness that this model shares with most overview documents is the absence of attempts to validate statistically the variables selected as probable determinants of site location. Of course, in cases where no existing data base is available or where the existing data are irretrievably flawed, this is the only possible approach. In other cases, however, there should be an effort to build a null, random model for the location of archaeological resources for statistical comparison with the actual distributions. Impressionistic isolation of determinant variables should be avoided since it may result in the use of variables whose significance cannot in fact be demonstrated or in the failure to use variables whose significance could be demonstrated. Even if the selected variables are the correct ones, the model will not be convincing to those who have other subjective impressions of the determinants of site location.

A Hierarchical Choice Model for Site Location

Before moving to predictive models based on deductive, optimizing assumptions and inductive models involving substantial analysis of ratio-level data, a brief discussion of an approach to settlement location analysis proposed by Limp and Carr (1985) will be useful. Their model should probably be categorized as a deductive approach to the systemic context, on an ordinal level of measurement (Figure 2.1). These authors propose that people make decisions about anything, including location of activities, by ranking the available alternatives into sets of equal preference value and then randomly selecting an alternative from the possibilities in the highest available preference set. This "general theory of rational choice"

was derived by Arrow (1951). The ordering of available options into these preference sets is based on "conditional preference aspects"—those aspects of the environment (broadly speaking) that directly bear on choices. When there is more than one "choice-making" aspect to be considered, it is assumed that the alternatives are evaluated in a sequential, hierarchical fashion by the decision-maker. In this framework an unfavorable aspect of a location (e.g., no water or too much water) cannot be mitigated by another, favorable aspect, as could happen in a linear additive model.

One key decision to be made in the analytic context when using this model is a choice as to how many preference sets should be assumed to have been in use for each choice-making aspect. If there are only two preference sets for each variable—satisfactory and unsatisfactory locations—the approach is formally identical to a "satisficing" approach (Simon 1957), as used by Williams et al. (1973) in the Great Basin, for example. As the number of sets that need to be ranked becomes greater than two for each variable, the framework approaches the optimization called for by classical marginalism: large numbers of bits of information have to be considered by both the decision-maker and the analyst. Intermediate numbers of preference sets imply an ordinal level of measurement.

Limp and Carr (1985) present a few brief examples of how this framework can be applied in different settings. They convincingly argue that hierarchical choice analysis is a realistic model for how people make decisions, since it does not assume that they can make, or wish to make, perfect calculation of return rates on every variable for every possible location. Nor are the data requirements in the analytic context as huge as for an optimal foraging theory model, for example. The hierarchical decision process assumed by this framework does not lend itself to easy discovery through any presently available computer algorithms, however, and it certainly cannot be reconstructed by such linear additive models as multiple linear regression, for example (Kohler and Parker 1986:428–430).

Generalizability. Because of its flexibility and its explicit reference to the systemic context, this model has very great generalizability. It has the ability to bring all kinds of choice-making aspects into consideration, not just those related to food resources. Indeed, one of the problems with the approach is that it is so very general that it gives few internal guidelines as to how it might be applied to a specific area. How many choice-making aspects should we expect? Where should the "break points" for a ratio-level variable like distance to water be established for each preference set, and how do we know this? Can an inferential technique be devised to reconstruct hierarchical decision frameworks from a distribution of points with and without archaeological resources? These are important questions that need to be addressed before application of such models can advance very far.

Besides these operational difficulties, we may ask to what extent it is appropriate to view all, or most, site locations as the result of "free" decisions in the systemic context. Kohler and Parker (1986:432–438) have identified a number of constraints on choice, instances in which "rational" decision rules are violated, cases where there is extreme lag in response to changing environmental determi-

nants, and other factors that make it difficult to analyze site location as though it were the outcome of simple, rational decisions. Then too, it will be suggested in Chapter 4 that settlement systems have a kind of internal logic that has little to do with individual or even group decisions at particular moments in time. Despite these very real problems, it is not easy to see how human behavior can be analyzed and predicted in the systemic context without considering how and why people make decisions.

Simplicity and *precision* cannot be evaluated for this example, since they depend on particular applications of the framework.

An Inferential Model for Site Location in Central and Southeastern Utah

The next case was selected as an example of the most common approach to predictive locational modeling in North America and particularly in the arid West. This is an inferential multivariate predictive model, operating on a ratio level of measurement and targeting the analytic context. This example, in common with many others that could be mentioned, is the result of a Class II (sample) cultural resource inventory—in this case, three tar sands areas in Utah (Schroedl 1984; Tipps 1984; Appendix, this volume).

For the larger two of the three study areas a two-phased random sample of quarter-sections was drawn, selecting 5 percent of the population on the first round and an additional 5 percent on the second round. (The third area was simply sampled at 10 percent, since it comprised only seven 160-acre quadrats.) The sequential samples were actually surveyed at the same time, but the results were recorded separately so that model building and model testing and revision could be conducted using different sets of data. Survey intensity and means of distinguishing sites and isolated finds are explicitly described in the report. For each site, probable age and cultural affiliation were recorded, and the site was classified into one of 10 descriptive site types (for example, pithouses, rockshelters, and lithic scatters with features). A second functional classification, more useful for explanatory purposes, was devised by evaluating eight criteria for the 158 sites/components in the sample:

1. diversity and size of the tool assemblage
2. maximum density of artifacts
3. frequency of debitage (lithic debris)
4. site size
5. number of features
6. type of features and amount of labor investment represented
7. presence of trash or midden deposits
8. presence of stratified deposits

The first five of these variables—those measured at the ratio level—were analyzed using principal components analysis (see Chapter 5). Four groups emerged on the

two significant factors, and these were interpreted as representing the major functional types suggested by Binford (1980) for logistically organized hunter-gatherers. The non-ratio-level variables were used to check the site classifications; these variables usually supported the type assignments made on the basis of the principal components analysis.

Site location analysis began with univariate descriptive frequencies for all sites in each study area with respect to elevation, aspect, slope, distance to permanent water, primary and secondary landform, depositional environment, primary and secondary vegetation, and primary and secondary geologic substrate.

One nice feature of this report is the discussion of how point estimates and confidence intervals for the total population of sites in each study area were calculated (Tipps 1984). It is relatively rare for confidence intervals to be calculated, which is a waste of one of the main advantages of random design adopted by most surveys. Tipps also warns her readers, quite correctly, that in two of three samples the amount of skewness relative to the sample size may lead to confidence intervals that are misleadingly narrow, using the normal parametric estimation techniques employed (for a discussion of statistical terms used here, see Chapter 5).

Two separate predictive models were developed (Schroedl 1984). One of these, incorporating Landsat imagery, turned out not to be very informative and will not be discussed further. Predictive models were constructed only for the two larger survey areas, which were somewhat more similar to each other than they were to the third area, and the two larger areas were pooled for purposes of analysis. Disappointingly, the functional identification of sites carefully worked out earlier in the report was not used for locational analysis and prediction, probably because of sample size considerations imposed by the inferential approach. (Division of the total pool of sites into its constituent classes significantly reduced the sample size in each class, which in turn makes it less likely that significant relationships with environmental variables will be discovered.) Nor is there any analysis of the location of the considerable number of isolated finds recorded during the survey.

The model-building process went through several preliminary stages. In the first, nine variables were used in a discriminant analysis to find the best linear function differentiating between sample quadrats from the initial 5 percent samples that contained, or did not contain, sites. Distances were measured from the center of each 160-acre quadrat. The directional aspect was broken into two components to avoid the problem typically associated with measurements in circular degrees. (A symptom of this problem is that 359° and 1° are very similar measurements.) The variables were

1. difference between the maximum and minimum elevation in each quadrat
2. distance to nearest permanent water
3. percentage of the quadrat covered by piñon-juniper
4. number of drainages within the quadrat
5. average quadrat elevation

6. distance to nearest river
7. distance to nearest wooded area
8. north-south aspect
9. east-west aspect

In the first two-group analysis, only the first four variables were selected by the stepwise procedure used for construction of the one discriminant function. Reclassification of the quadrats on which the function was based into their original groups (sites, no sites) was 73 percent successful; classification error rates for the quadrats of the second 5 percent sample were about 10 percent higher. These results are somewhat lower than, although within the range of, other similar attempts tabulated by Schroedl (1984:155). A second stage of refinement, which involved discarding three outliers from the analysis and using more of the sample quadrats in the initial classification-building portion of the discriminant analysis, improved these results; two additional variables (5 and 6 above) also contributed to the linear discriminant function.

The final analysis employed all of the sample quadrats and discriminated three groups of quadrats: those without sites, those with one site, and those with more than one site. Reclassification rates were quite high but, of necessity, were based on the same sample for which the functions were obtained in the first place. In a three-group solution there may be one or two significant linear discriminant functions; there are two in this example. The first, explaining about 40 percent of the total variance, showed that high-elevation quadrats with relatively large proportions of piñon-juniper contained a larger number of sites than low-elevation, unwooded quadrats. The second function, which explained about 12 percent of the total variance, was orthogonal to the first; that is, this function exploited a dimension of variability uncorrelated with the high elevation/high piñon-juniper vs low elevation/low piñon-juniper dimension. Apparently there were several quadrats that had a relatively high number of drainages but were not significantly higher in elevation than those having only a few drainages. These same quadrats were also located a long way from a river and tended to contain only one site; they were differentiated from quadrats with no sites or with two or more sites along this dimension.

Generalizability. It seems probable that this solution exploits a good deal of variability peculiar to this particular sample; it would be surprising if the second dimension of variability turned out to be typical of much of the intermountain West. The first dimension is much more general; a similar discriminator could probably be found in many areas at similar elevations in the intermontane region.

Simplicity. The final predictive model, in the form of the classificatory equations derived from the discriminant analysis, allows unambiguous classification of any quadrat from the spatial population into one of the three groups on the basis of measurements on six variables (the original nine variables less distance to wooded area and the two aspect determinations).

Precision. The 160-acre quadrats do not allow for very precise prediction of site location. The author points out, quite reasonably, that achieving higher spatial precision over large areas is extremely time consuming without the use of such computerized data-collection aids as geographic information systems (see Chapter 10). Nor are the predictions very fine grained in terms of the types of sites predicted to be present. Some gain in precision in terms of the number of sites predicted for unsurveyed quadrats is achieved by the three-group solution, in contrast to the a priori site/nonsite classes used by most analysts. There is little reason to expect, however, that the local environment in quadrats with one site should be opposable to that in quadrats with more than one site along a continuum that is at right angles (or uncorrelated with) the continuum that distinguishes between quadrats with no sites and quadrats with many sites. Some functional differentiation in site types is almost certainly being exploited here, and the results might have been even better had this distinction been taken into account for prediction.

An Optimal Foraging Theory Model of Site Location for the Northeastern Continental Shelf

Barber and Roberts (1979) present both an inductive and a deductive approach to the difficult problem of estimating site types and densities on those portions of the continental shelf from the Bay of Fundy in Maine to Cape Hatteras in North Carolina that are now submerged but were exposed at or after 18,000 BP. Although they face unusual measurement problems because of the nature of their study area, their conceptualization problems are the same as those for a dry-land model. Only their deductive model—based on optimal foraging theory—will be discussed here; see the Appendix for a summary of the entire project.

Optimal foraging theory models are derived from fundamental assumptions in evolutionary ecology and population genetics in which change in the relative frequency of traits in a population is interpreted as being due to differential inclusive fitness among the individuals in that population. From this perspective, the goal of behavior should be to maximize the individual's proportionate contribution to the genotype of the next generation. Unfortunately, inclusive fitness is difficult or impossible to measure, but it may have correlates that can be measured. Optimal foraging theory assumes that the net rate of energy captured by an individual (or some similar measure) is such a correlate, and that it will be maximized by selective forces (Smith 1980:58).

There has been an extended discussion about the applicability of such models to human populations. Those cultural ecologists who accept the "monism" dictate discussed above consider these models to be clearly relevant. Eric Alden Smith (1980:12–15) points out that there is a middle ground between two extreme positions: *(a)* that cultural processes are perfectly analyzable in terms of general evolutionary models, with the only meaningful distinction being that cultural evolution is more rapid and more finely tuned; and *(b)* that cultural processes are shaped by purely cultural goals that have no necessary congruence with biological criteria for adaptation.

> [A] third alternative is . . . that cultural criteria guiding decisions and long-term changes are closely correlated with, but not isomorphic with, biological criteria of adaptive success. In this case, biological factors acting at the proximate level ensure that cultural modes of inheritance will not substitute selective criteria that are consistently in conflict with fitness maximization. . . .
>
> Selective criteria of genetic evolution, and those of culture change or individual decision-making, will be generally but not perfectly correlated [1980:14–15].

Models for the location of behavior based on optimal foraging theory share some similarities with the general choice theory used by Limp and Carr. Since they deal in decisions, both operate within the systemic context. The hierarchical choice methods essentially specify *how* choices are made (choice mechanisms), however, while optimal foraging theory also specifies *why* choices are made (choice goals). In one sense the approach advocated by Limp and Carr is more generalizable, since goals other than optimizing food intake can be accommodated. Optimal foraging theory is more complete, and perhaps more useful, however, since it contains internal guidelines to predict exactly what choices will be made given an array of information on resource costs. Both use a deductive logic for prediction.

The information needed to apply and test optimal foraging models is difficult and expensive to collect, and it has not been easy to test such models, even in modern ethnographic contexts (but see Smith 1980; Winterhalder 1983). In the archaeological context the problems are multiplied immensely. These problems are particularly serious for the application discussed here, since no detailed paleoenvironmental maps are available for the inundated continental shelf. For some of the resources, return rates have been experimentally estimated by Perlman (1976). Since the rigorous quantification of net resource yields called for by optimal foraging theory was impossible for most resources, the authors dichotomize the major potential food resources along two dimensions: the probable importance of the resource, based on grossly estimated caloric return rates (primary vs secondary resources), and the degree to which location in the immediate vicinity of the resource is necessary for efficient exploitation of that resource (determinate vs indeterminate resources). Shellfish, for example, have relatively low return rates and are therefore secondary, but they are localized in space and have a large amount of waste weight, which would encourage location of sites in the vicinity of the resource (Barber and Roberts 1979:306). The resources characterized in this manner are shown in Table 2.1.

The authors recognize that the immediate predictions made by optimal foraging theory concern what resource patches will be exploited under what conditions. Locations of settlements, therefore, are one order of inference removed from the predictions that optimal foraging theory is designed to make. The spatial resolution of predictions is so low for this particular model, however, that this may not be a problem. Only four zones are differentiated for prediction: full coastal, estuarine, inland valley, and upland. Given that the locations of these zones change during marine transgression, Barber and Roberts separate the extremely long period of interest (beginning at 18,000 BP) into six 3000-year segments. They also

TABLE 2.1.

Resources categorized by return rate and role in influencing site location

Resource	*Importance (Return Rate)*	*Role in Determining Location*
nuts (various oaks, hickory, and pecan)	primary	indeterminate
mammoth and mastodon	secondary	indeterminate
caribou	secondary	indeterminate
moose	secondary	indeterminate
seals (spring)	primary	determinate
(other seasons)	secondary	indeterminate
walrus	secondary	indeterminate (?)
anadromous fish	primary	determinate
other fish	primary (?)	determinate (?)
marine molluscs	secondary	determinate

From Barber and Roberts 1979:307-314

subdivide the north-south expanse of continental shelf into three subareas: Maine, southern New England, and Mid-Atlantic.

For each period, in each subarea, predictions are made concerning the probable site size, site density, and to a limited extent, site type in each of the four environmental zones (a portion of one of their tables is reproduced here as Table 2.2). The authors assume that site size is correlated with population size; dispersed populations will be found in areas with "predictable, mobile, and evenly distributed resources," leading to small sites. Aggregated populations and, consequently, large sites will be found in areas with unpredictable, immobile, and clumped resources (Barber and Roberts 1979:316). The effects on site size of such variables as duration of occupation and location reuse are not considered. Site density, in turn, is considered to be a function of the "relative attractiveness of the several environments for exploitation" (1979:317) and so is predicted only on an ordinal level within each period, for each subarea. Barber and Roberts intend these projections of site size and frequency to be suggestive; they do not believe that more precise estimates could be calculated reliably using available information.

Generalizability and Precision. Models in which both decision mechanisms and decision goals are fully specified by theory seem to provide the only consistently deductive, truly rigorous formulation for predicting site location. For optimal foraging theory models the resources actually used must be inferred for each specific application, and return rates for these resources must be calculated for each case. Once these inferences and calculations have been made, however, all predictions as to resource use then follow automatically from the theory itself. This is in contrast to the rational choice theory approach described above, or to the satisficing approach, where preference sets or acceptability criteria must also be discovered inferentially or made up using rules of thumb.

When optimal foraging theory models are used to predict the locations of activities resulting in the deposition of archaeological materials, the explicit focus

TABLE 2.2.

Example predictions

Subarea	*Time Span (BP)*	*Paleoenvironment*	*Predicted Site Size*	*Predicted Site Density*	*Shell Middens Present*	*Anadromous-Fish Camps Present*
Maine	18,000-12,000	under glacier or sea	none	none		
	12,000-9000	full coastal	small	low		
		estuarine	small	low		
		inland valley	very small	low		X
		upland	very small	low		
	9000-6000	full coastal	small-medium	medium-low	X	
		estuarine	small	medium	X	X
		inland valley	small	low		X
		upland	small	low		
	6000-3000	full coastal	small-large	medium-high	X	
		estuarine	small-medium	medium	X	X
		inland valley	small-medium	medium		X
		upland	small	low		

From Barber and Roberts 1979:322

on the spatial distribution and return rates of food resources (only) is a two-edged sword. There is no reason to doubt that there will be a general correlation between the distributions of archaeological materials and the distribution of exploited resources; after two decades of settlement pattern analyses this is no longer a surprising conclusion, or even one worthy of research in itself. Considerably more work *is* needed, however, on predicting exactly what these materials will be, how they were deposited, and what their relationship was to other materials elsewhere on the landscape. This task will require consideration of more than the distribution of food resources.

Optimal foraging theory assumes that all humans are foragers. In Chapter 4 we will argue that, since not all humans are foragers, the degree of intensification affects the organization of the settlement systems, and this in turn determines how spatially predictable the sites generated by that system will be on the basis of variables in the natural environment alone. For example, in the case of foragers we might expect that many resource patches—especially if they overlap spatially in their temporal availability with other nonsubstitutable resources and are relatively isolated rather than continuous in their spatial distribution—will support residential bases. These same resource patches, however, may be visited intermittently by specialized task groups in a logistically organized subsistence system. In still more intensified systems, variables other than the distribution of environmental resources become increasingly important in the location of residences and other site types. We need to begin trying to make predictions with more specificity about how human settlement systems interact with the environment—not just where undifferentiated sites or materials will end up on the landscape, but what kind of use these represent in the systemic context.

The lack of behavioral (and spatial) precision is no fault of these particular authors, who suffered more severe measurement problems than most. No large predictive locational models have ever been constructed with great behavioral specificity. These considerations are relevant here, I believe, because if such specificity is ever to be achieved it will be through a deductive approach to the systemic context, using detailed reconstructions of the resource availability in the paleoenvironment.

The generalizability of optimal foraging theory models for human use of the landscape is limited by their relatively low degree of "portability" across different adaptation types, especially those of increasing intensification. The precision with which these models do what they were explicitly designed to do—predict foraging exploitation of resource patches—is probably high in the ideal case, although this is difficult to test. In the particular example discussed here, measurement problems interfere with high precision.

Simplicity. Optimal foraging theory models are wonderful in the simplicity of their design and the economy of their assumptions. In fact, it is this very simplicity that prevents them from being more general. What is not simple, however, is handling the mass of ratio-level information necessary to rigorously map a predic-

tive model based on this theory. For such purposes a geographic information system (Chapter 10) seems essential.

I do not mean to minimize the shortcomings of optimal foraging theory, particularly as they might affect the accuracy of prediction. One such shortcoming is the assumption that each resource patch, and the landscape as a whole, will be used at its maximum capacity, when in fact hunter-gatherers typically do not expand their populations to the carrying capacity of the region. Another is that cultures frequently have high-status resources (and conversely exhibit food taboos) that do not have any obvious relation to resource abundance or caloric content. Readers should consult Martin (1983, 1985), Sih and Milton (1985), Hawkes and O'Connell (1985), Yesner (1985), and Smith and Winterhalder (1985) to capture the complexity of the recent debate on issues surrounding application of optimal foraging theory to human societies.

Discussion

Generalizability. One clear conclusion emerges from these four examples: deductive theories of settlement location that work from first principles have considerably more potential generalizability than do specific models designed for particular areas and derived almost entirely through empirical procedures. Thus, the framework of decision theory and analysis discussed by Limp and Carr (1985) is very generalizable; the optimal foraging theory framework is somewhat less generalizable but can still be applied to differing environments and adaptation types.

The inductive or inferential framework, *as an overall strategy,* is very generalizable. An inductive model can be constructed for any area that has a partially known archaeological or ethnographic record. But we must differentiate between a strategy for analysis or prediction (inductive generalization vs deductive implication) and a model explaining or predicting site location. An optimal foraging theory model can be applied in any area; only the structure of the environment in question, and the resources actually used, change. Each new inferential model starts from scratch: of the infinity of variables that might have affected how people used space, which actually did?

Precision. There is no inherent difference between inferential and deductive models in their potential spatial resolution of prediction. As it happens, none of the models discussed above had finely resolved spatial predictions, although some inferential models (for example, those discussed by Kvamme later in this volume) do. There is more to precision than spatial resolution, however. How fine-grained are the predictions of site type or of the cultural and natural forces at work in the formation of the archaeological record? Are predictions made about assemblage content? As such questions approach behavior in the systemic context more closely, it becomes more natural to frame them in a deductive manner.

Simplicity. The discussion of the two deductive models for site location suggests that there is a general trade-off between simplicity and generalizability.

The optimal foraging theory model is more parsimonious but less generalizable than hierarchical choice theory. Inferentially constructed models are not necessarily more parsimonious than deductive models. Although the examples used here shed no light on this question, Limp and Carr (1985) suggest that a few processes can generate a multiplicity of forms. Since inferential models deal with forms, and deductive models with processes, the latter may prove more parsimonious.

CONCLUSIONS

Some of what has been said above seems to favor deductive approaches over inferential approaches to the problem of predicting types and locations of archaeological materials, and the same will be true of the method and theory discussion in Chapter 4. And yet, while models are classified one way or another here for taxonomic purposes, it is evident that neither purely deductive nor purely inductive models are possible. In the first case, we would not know how to apply the model to a particular area; in the second, we would not know what variables should be considered for inclusion in the analysis.

Much of this book will be devoted to discussing the kinds of inductively derived models that constitute most current efforts in archaeological predictive modeling. While empirical correlative models can be very useful in specific cases, in this chapter and in Chapter 4 we would like to balance the picture somewhat by suggesting that deductive explanatory models should have greater utility in the long run. Both the manager and the researcher want predictive models that are *useful,* after all, and as Blalock (1979:120) points out, there are several ways that utility can be defined in such a context. One of these is in the significance of what we learn through the application of the model. I think that nearly everyone will agree that it is more significant to learn something about *both* the systemic and archaeologic contexts at the same time than it is to learn about the archaeologic or analytic context alone, as is so often the case for inferential models.

Another indication of utility is whether the application of the model results in predictions that go beyond those that could have been made by common sense or by a casual examination of the phenomena in question. As long as we couch our analyses in terms of casually observable variables (for example, a dichotomy between site presence or absence) it will be hard to transcend common sense predictions, such as the prediction that sites will cluster around resources basic to human needs.

A third potential criterion for utility is the generalizability of a model to other times and places. In fact, until such time as we begin to gather reliable estimates of model accuracy, I suggest that we strive to build models that are both generalizable and precise. If generalizable and precise models can be constructed, I think we will find that accuracy will take care of itself.

I would like to acknowledge the support of the School of American Research in Santa Fe, where I was residing and working when I made the changes in the text in response to the review comments.

REFERENCES CITED

Ammerman, Albert J.
1981 Surveys and Archaeological Research. *Annual Reviews in Anthropology* 10:63–88.

Arrow, K.
1951 *Social Choice and Individual Values.* Yale University Press, New Haven.

Barber, Russell, and Michael E. Roberts
1979 *Archaeology and Paleontology.* A Summary and Analysis of Cultural Resource Information on the Continental Shelf from the Bay of Fundy to Cape Hatteras, Final Report, vol. II. Institute for Conservation Archaeology Report 88. Peabody Museum, Harvard University, Cambridge.

Bates, Marston
1953 Human Ecology. In *Anthropology Today: An Encyclopedic Inventory,* edited by A. L. Kroeber, pp. 700–713. University of Chicago Press, Chicago.

Bennett, J. W.
1946 An Interpretation of the Scope and Implications of Social Scientific Research in Human Subsistence. *American Anthropologist* 48:553–573.

Bettinger, Robert L.
1980 Explanatory/Predictive Models of Hunter-Gatherer Adaptation. In *Advances in Archaeological Method and Theory,* vol. 3, edited by Michael B. Schiffer, pp. 189–255. Academic Press, New York.

Binford, Lewis R.
1980 Willow Smoke and Dogs' Tails: Hunter-Gatherer Settlement Systems and Archaeological Site Formation. *American Antiquity* 45:1–17.

Blalock, Hubert M., Jr.
1979 Dilemmas and Strategies of Theory Construction. In *Contemporary Issues in Theory and Research: A Metasociological Perspective,* edited by William E. Snizek, Ellsworth R. Fuhrman, and Michael K. Miller, pp. 119–135. Contributions in Sociology 33. Greenwood Press, Westport, Connecticut.

1982 *Conceptualization and Measurement in the Social Sciences.* Sage Publications, Beverly Hills.

Boyce, Ronald R., editor
1980 *Geography as Spatial Interaction.* University of Washington Press, Seattle.

Brookfield, Harold
1968 New Directions in the Study of Agricultural Systems in Tropical Areas. In *Evolution and Environment: A Symposium Presented on the Occasion of the One-Hundredth Anniversary of the Foundation of the Peabody Museum of Natural History at Tale University,* edited by Ellen T. Drake, pp. 413–439. Yale University Press, New Haven.

1969 On the Environment as Perceived. *Progress in Geography: International Review of Current Research* 1:51–80.

Burgess, R. J., K. L. Kvamme, P.R. Nickens, A. D. Reed, and G.C. Tucker, Jr.
1980 *Class II Cultural Resource Inventory of the Glenwood Springs Resource Area, Grand Junction District, Colorado.* Nickens and Associates. Submitted to the Bureau of Land Management, Contract No. YA-512-CT9-222. Copies available from the Bureau of Land Management, Grand Junction District, Grand Junction, Colorado.

Clark, J. G. D.
1952 *Prehistoric Europe: The Economic Basis.* Methuen, London.

Clarke, David L.
1968 *Analytical Archaeology.* Methuen, London.

Conklin, Harold
1961 The Study of Shifting Cultivation. *Current Anthropology* 2:27–61.

Damas, David
1969 The Study of Cultural Ecology and the Ecology Conference. In *Contributions to Anthropology: Ecological Essays*, edited by David Damas, pp. 1–12. National Museums of Canada Bulletin 230.

Darsie, Richard F., III
1983 *Prehistoric Agricultural Potential and Site Location in the Dolores River Valley, Southwest Colorado.* Unpublished M.A. Thesis, Department of Anthropology, Washington State University, Pullman.

Dean, Jeffrey S.
1978 An Evaluation of the Initial SARG Research Design. In *Investigations of the Southwestern Anthropological Research Group: The Proceedings of the 1976 Conference*, edited by Robert C. Euler and George J. Gumerman, pp. 103–117. Museum of Northern Arizona, Flagstaff.

Dincauze, Dena F. and Judith W. Meyer
1976 *Prehistoric Resources of East-Central New England: A Preliminary Predictive Study.* University of Massachusetts, Amherst. Submitted to the National Park Service, Contract No. 0001-4-068. NTIS No. PB-265-019.

Ellen, Roy
1982 *Environment, Subsistence and System: The Ecology of Small-Scale Formations.* Cambridge University Press, Cambridge, England.

Evans-Pritchard, Sir Edward
1981 *A History of Anthropological Thought.* Basic Books, New York.

Flannery, Kent V.
1968 Archeological Systems Theory and Early Mesoamerica. In *Anthropological Archaeology in the Americas*, edited by Betty J. Meggers, pp. 67–87. Anthropological Society of Washington, Washington, D.C.

Fuller, Steven., A. E. Rogge, and Linda M. Gregonis
1976 *Orme Alternatives: The Archaeological Resources of Roosevelt Lake and Horseshoe Reservoir.* 2 vols. Archaeological Series 98. Arizona State Museum, Tucson.

Glacken, Clarence J.
1967 *Traces on the Rhodian Shore: Nature and Culture in Western Thought from Ancient Times to the End of the Eighteenth Century.* University of California Press, Berkeley.

Gold, John R.
1980 *An Introduction to Behavioural Geography.* Oxford University Press, Oxford.

Green, Ernestine L.
1973 Location Analysis of Prehistoric Maya Sites in Northern British Honduras. *American Antiquity* 38:279–293.

Green, Stanton
1980 Broadening Least-Cost Models for Expanding Agricultural Systems. In *Modeling Change in Prehistoric Subsistence Economies*, edited by T. K. Earle and A. L. Christenson, pp. 209–241. Academic Press, New York.

Hackenberger, Steven
1978 Reconnaissance Sampling of Fort Sill: Methods and Results. In *An Archaeological Reconnaissance of Fort Sill, Oklahoma*, assembled by C. Reid Ferring, pp. 118–161. Contributions of the Museum of the Great Plains 6. Lawton, Oklahoma.

Haggett, P., A. D. Cliff, and A. Frey
1977 *Locational Analysis in Human Geography*, 2nd ed. Wiley, New York.

Hardesty, D. L.
1975 The Niche Concept: Suggestions for its Use in Studies of Human Ecology. *Human Ecology* 3:71–85.

Harris, Marvin
1968 *The Rise of Anthropological Theory: A History of Theories of Culture.* Thomas Crowell, New York.

Haury, Emil W.
1956 Speculations on Prehistoric Land Settlement Patterns in the Southwest. In *Prehistoric Settlement Patterns in the New World*, edited by Gorden R. Willey, pp. 3–10. Viking Fund Publications in Anthropology 23. New York.

Hawkes, Kristen, and James F. O'Connell
1985 Optimal Foraging Models and the Case of the !Kung. *American Anthropologist* 87:401–405.

Heizer, Robert F., and Martin A. Baumhoff
1956 California Settlement Patterns. In *Prehistoric Settlement Patterns in the New World*, edited by Gorden R. Willy, pp. 32–44. Viking Fund Publications in Anthropology 23. New York.

Hill, James N.
1971 Research Propositions for Consideration, Southwestern Anthropological Research Group. In *The Distribution of Prehistoric Population Aggregates*, edited by George J. Gumerman, pp. 38–44. Prescott College Press, Prescott, Arizona.

Hodgen, Margaret T.
1964 *Early Anthropology in the Sixteenth and Seventeenth Centuries.* University of Pennsylvania Press, Philadelphia.

Holmer, Richard N.
1979 *Split Mountain Cultural Study Tract.* Archeological Center, Department of Utah. Submitted to the Bureau of Land Management, Contract No. UT-080-CT-9-003. Copies available from the Bureau of Land Management, Vernal District, Vernal, Utah.

Interagency Archeological Services (IAS)
1976 *A Status Report to the Archeological Community.* National Park Service, USDI, Washington, D.C.

Jochim, Michael A. 1976 *Hunter-Gatherer Subsistence and Settlement: A Predictive Model.* Academic Press, New York.

King, Thomas F.
1978 *The Archeological Survey: Methods and Uses.* Cultural Resources Management Studies. Office of Archeology and Historic Preservation, Heritage, Conservation, and Recreation Service, Washington, D.C.

1984 An Overview of Archeological Law at the Federal Level. *American Archeology* 4:115–118.

Kohl, P. J.
1981 Materialist Approaches in Prehistory. *Annual Reviews in Anthropology* 10:89–118.

Kohler, Timothy A.
1986 Grass Mesa Archaeological Locality Overview, 1979. In *Dolores Archaeological Program: Anasazi Communities at Dolores: Early Small Settlements in the Dolores River Canyon and Western Sagehen Flats Area*, compiled by Timothy A. Kohler, William D. Lipe, and Allen E. Kane, pp. 17–44. Bureau of Reclamation, Denver.

Kohler, Timothy A., and Sandra C. Parker
1986 Predictive Models for Archaeological Resource Location. In *Advances in Archaeological Method and Theory*, vol. 9, edited by Michael B. Schiffer, pp. 397–452. Academic Press, New York.

Kohler, Timothy A., Thomas P. DesJeans, Carl Feiss, and Don F. Thompson
1980 *An Archaeological Survey of Selected Areas of the Fort Benning Military Reservation, Alabama and Georgia.* Remote Sensing Analysts. Submitted to the United States Army, Contract No. C-5716(78). Copies available from the Interagency Archeological Services, Atlanta.

Kohler, Timothy A., Carl J. Phagan, and Eric Blinman
1985 Sources of Confounding Variability in Archaeological Collections. Dolores Archaeological Program Technical Reports DAP-282. Draft submitted to the Bureau of Reclamation, Contract No. 8-07-40-50562. Copies available from the Bureau of Reclamation, Upper Colorado Region, Salt Lake City.

Kroeber, A. L.
1939 *Cultural and Natural Areas of Native North America.* University of California Publications in American Archaeology and Ethnology 38. Berkeley.

Kvamme, Kenneth
1980 Predictive Model of Site Location in the Glenwood Springs Resource Area. In *Class II Cultural Resource Inventory of the Glenwood Springs Resource Area, Grand Junction District, Colorado.* Submitted to the Bureau of Land Management, Grand Junction District Office, Colorado.

Limp, W. Frederick, and Christopher Carr
1985 The Analysis of Decision Making: Alternative Applications in Archeology. In *The Analysis of Archeological Data Structures*, edited by Christopher Carr. Institute for Quantitative Archeology, Fayetteville, Arkansas, and Mayport Press, Kansas City, in press.

Marsh, George P.
1864 *Man and Nature, or Physical Geography as Modified by Human Action.* Charles Scribner, New York.

Martin, John F.
1983 Optimal Foraging Theory: A Review of Some Models and Their Applications. *American Anthropologist* 85:612–629.

1985 More on Optimal Foraging Theory. *American Anthropologist* 87:649–650.

Mierendorf, Robert R., Terry Keller, Denise Carlevato, and Patricia A. McLeod
1981 *A Cultural Resources Predictive Land Use Model for the Okanogan Highlands: An Impact Assessment for the Bonneville Power Administration's Chief Joseph Dam–Tonasket and Grand Coulee–Keller Transmission Lines.* Bonneville Cultural Resources Group Report 100-2. Eastern Washington University Reports in Archaeology and History. Cheney, Washington.

Muto, Guy R., and Joel Gunn
1980 A Study of Late Quaternary Environments and Early Man along the Tombigbee River, Alabama and Mississippi. Phase I Interim Report. Benham Blair and Affiliates. Draft submitted to the Heritage, Conservation, and Recreation Service, Washington, D.C.; Interagency Archeological Services, Atlanta; and Corps of Engineers, South Atlantic District, Mobile. Copies available from Benham Blair and Affiliates, Oklahoma City.

Nance, C. Roger, Harry Holstein, and David C. Hurst
1983 Evaluation of Multiple Regression Models Predicting Archaeological Site Distributions at Fort McClellan, Alabama. Paper presented at the 48th Annual Meeting of the Society for American Archaeology, Pittsburgh.

Netting, R. McC.
1974 Agrarian Ecology. *Annual Reviews in Anthropology* 3:21–56.

Parsons, Jeffrey R.
1972 Archaeological Settlement Patterns. *Annual Reviews in Anthropology* 1:127–150.

Paynter, Robert
1982 *Models of Spatial Inequality: Settlement Patterns in Historical Archaeology.* Academic Press, New York.

Perlman, Stephen M.
1976 *Optimum Diet Models and Prehistoric Hunter-Gathers: A Test Case on Martha's Vineyard.* Unpublished Ph.D. dissertation, Department of Anthropology, University of Massachusetts, Amherst.

Pilgram, Thomas Kurt
1982 *Predicting Archaeological Sites from Environmental Variables: A Mathematical Model for the Sierra Nevada Foothills.* Ph.D dissertation, University of California, Berkeley. University Microfilms, Ann Arbor.

Plog, Fred T., and James N. Hill
1971 Explaining Variability in the Distribution of Sites. In *The Distribution of Prehistoric Population Aggregates,* edited by George J. Gumerman, pp. 7–36. Prescott College Anthropological Reports 1. Prescott College Press, Prescott, Arizona.

Rappaport, Roy A.
1971 The Flow of Energy in an Agricultural Society. *Scientific American* 225:117–132.

Ratzel, Friedrich
1896–1898 *Anthropogeographie,* translated by A. J. Butler from the second edition. MacMillan, New York.

Robertson, Ben P., and Linda Butler Robertson
1978 The Generation of Locational Models in an Inductive Framework. In *Conservation Archaeology in the Northeast: Toward a Research Orientation,* edited by Author E. Spiess, pp. 27–36. Peabody Museum Bulletin 3. Peabody Museum, Harvard University, Cambridge.

Roper, Donna C.
1979 *Archaeological Survey and Settlement Pattern Models in Central Illinois.* Illinois State Museum Scientific Papers, vol. XVI. Midcontinental Journal of Archaeology Special Papers No. 2.

Rosenberg, Ronald H.
1984 Archaeological Resource Preservation: The Role of State and Local Government. *American Archeology* 4:125–131.

Sanders, William T., Jeffrey R. Parsons, and Robert S. Santley
1979 *The Basin of Mexico: Ecological Processes in the Evolution of a Civilization.* Academic Press, New York.

Schiffer, Michael B.
1972 Archaeological Context and Systemic Context. *American Antiquity* 37:156–165.

Schroedl, Alan R.
1984 Site Location and Predictive Modelling. In *The Tar Sands Project: Cultural Resource Inventory and Predictive Modelling in Central and Southern Utah,* edited by Betsy L. Tipps, pp. 141–169. P-III Associates. Submitted to the Bureau of Land Management, Contract No. YA551-CT3-340038. Copies available from Bureau of Land Management, Richfield, Utah.

Scott, Thomas R., Michael McCarthy, and Mark A. Grady
1978 *Archaeological Survey in Cherokee, Smith and Rusk Counties, Texas: A Lesson in Survey Methods.* Archaeological Research Report 116. Department of Anthropology, Southern Methodist University, Dallas.

Scovill, Douglas H.
1974 History of Archaeological Conservation Policy and the Moss-Bennett Bill. In *Proceedings of the 1974 Cultural Resource Management Conference, Federal Center, Denver, Colorado,* edited by William D. Lipe and Alexander J. Lindsay, Jr., pp. 1–11. Technical Series 14. Museum of Northern Arizona, Flagstaff.

Sears, William H.
1956 Settlement Patterns in the Eastern United States. In *Prehistoric Settlement Patterns in the New World*, edited by Gordon R. Willey, pp. 45–51. Viking Fund Publications in Anthropology 23. New York.

Semple, Ellen C.
1911 *Influences of Geographic Environment on the Basis of Ratzel's System of Anthropo-Geography*. Henry Holt, New York.

Senour, Kenneth, Jr., compiler
1980 Field/Technical Reports Supporting the Bull Creek Environmental Impact Statement, White County, Arkansas, vol. 3. Technology Research and Development, Inc. Draft submitted to the Mid-Arkansas Regional Water Distribution District. Copies available from Mid-Arkansas Regional Water Distribution District, Oklahoma City.

Sih, Andrew, and Katharine A. Milton
1985 Optimal Diet Theory: Should the !Kung Eat Mongongos? *American Anthropologist* 87:395–401.

Simon, H. A.
1957 *Models of Man: Social and Rational*. Wiley, New York.

Smith, C. A., editor
1976 *Regional Analysis*, 2 vols. Academic Press, New York.

Smith, Eric Alden
1980 *Evolutionary Ecology and the Analysis of Human Foraging Behavior: An Inuit example from the East Coast of Hudson Bay*. Ph.D. dissertation, Cornell University. University Microfilms, Ann Arbor.

Smith, Eric Alden, and Bruce Winterhalder
1985 On the Logic and Application of Optimal Foraging Theory: A Brief Reply to Martin. *American Anthropologist* 87:645–648.

Steward, Julian
1938 *Basin-Plateau Aboriginal Sociopolitical Groups*. Bureau of American Ethnology Bulletin 120. Washington, D.C.
1955 *The Theory of Culture Change*. University of Illinois Press, Urbana.
1973 *Alfred Kroeber*. Columbia University Press, New York.

Stewart, Omer C.
1943 Notes on Pomo Ethnogeography. *University of California Publications in American Archaeology and Ethnology* 40(2):29–62. Berkeley.

Sullivan, Alan P., and Michael B. Schiffer
1978 A Critical Examination of SARG. In *Investigations of the Southwestern Anthropological Research Group: The Proceedings of the 1976 Conference*, edited by Robert C. Euler and George J. Gumerman, pp. 168–175. Museum of Northern Arizona, Flagstaff.

Thomas, David Hurst
1973 An Empirical Test for Steward's Model of Great Basin Settlement Patterns. *American Antiquity* 38:155–176.

Tipps, Betsy L.
1984 Site Density Estimation. In *The Tar Sands Project: Cultural Resource Inventory and Predictive Modelling in Central and Southern Utah*, edited by Betsy L. Tipps, pp. 127–140. P-III Associates. Submitted to the Bureau of Land Management, Contract No. YA551-CT3-340038. Copies available from the Bureau of Land Management, Richfield, Utah.

Trigger, Bruce G.
1968 The Determinants of Settlement Patterns. In *Settlement Archaeology*, edited by K.C. Chang, pp. 53–78. National Press Books, Palo Alto.
1971 Archaeology and Ecology. *World Archaeology* 2:321–336.

Ullman, Edward L.
1941 A Theory of Location for Cities. *American Journal of Sociology* 41:853–864.

Vita-Finzi, C., and E. S. Higgs
1970 Prehistoric Economy in the Mount Carmel Area of Palestine: Site Catchment Analysis. *Proceedings of the Prehistoric Society* 36:1–37. Cambridge, England.

Weber, A.
1929 *Theory of the Location of Industries*, translated by C. J. Friedrich. University of Chicago Press, Chicago.

White, Leslie A.
1949 *The Science of Culture.* Farrar, Straus, New York.

Wildesen, Leslie E.
1974 Archaeologists and Planners: The Uses and Misuses of Predictive Models. Paper presented at the 1974 Annual Meeting of the Society for California Archaeology, Riverside.

Willey, Gordon R.
1953 *Prehistoric Settlement Patterns in the Viru⅛* Valley, Peru. Bureau of American Ethnology Bulletin 155. Washington, D.C.

Willey, Gordon R., editor
1956 *Prehistoric Settlement Patterns in the New World.* Viking Fund Publications in Anthropology 23. New York.

Williams, L., D. H. Thomas, and R. L. Bettinger
1973 Notions to Numbers: Great Basin Settlements as Polythetic Sets. In *Research and Theory in Current Archaeology*, edited by C. L. Redman, pp. 215–237. Wiley, New York.

Williams, Stephen
1956 Settlement Patterns in the Lower Mississippi Valley. In *Prehistoric Settlement Patterns in the New World*, edited by Gordon R. Willey, pp. 52–62. Viking Fund Publications in Anthropology 23. New York.

Winterhalder, Bruce
1981 Optimal Foraging Strategies and Hunter-Gatherer Research in Anthropology: Theories and Models. In *Hunter-Gatherer Foraging Strategies*, edited by B. Winterhalder and E. A. Smith, pp. 66–98. University of Chicago Press, Chicago.

1983 Boreal Foraging Strategies. In *Boreal Forest Adaptations*, edited by A. Theodore Steegmann, Jr., pp. 201–241. Plenum Press, New York.

Wissler, Clark
1922 *The American Indian: An Introduction to the Anthropology of the New World.* Oxford University Press, New York.

Woodward-Clyde Consultants
1978 Appendix E. In Well Field Development Plan for the Wasson Field-Denver Unit CO_2 Project Environmental Impact Report. Woodward-Clyde Consultants. Draft submitted to Shell Oil Corporation. Copies available from Woodward-Clyde Consultants, Three Embarcadero Center, Suite 700, San Francisco CA 94111.

Yesner, David R.
1985 Archeology and Optimal Foraging Theory: Appropriate Analytical Units. *American Anthropologist* 87:412–414.

Chapter 3

MODELS AND THE MODELING PROCESS

Jeffrey H. Altschul

In Chapter 2 a model was defined as "a simplified set of testable hypotheses." Researchers investigating a particular phenomenon create a model by isolating various components of the phenomenon and then positing (or hypothesizing) a relationship or series of relationships among them. The result is a simplified version of the phenomenon that mimics, in a general way, the events or behaviors in question.

One of the utilities of a model is that it is possible to hypothesize how changes in one or more components will affect the final state of the phenomenon; that is, one can predict what the phenomenon will "look like" given specified changes in particular components. All models are predictive in this sense. It is important to emphasize, however, that prediction is not synonymous with explanation and that predictive accuracy alone is not necessarily the best indicator of a model's utility. For instance, the old adage

> Red sky at night, sailor's delight;
> Red sky in morning, sailor take warning

is a perfectly valid predictive model of the weather. Based on a single observation one can predict whether or not there will be a storm in the immediate future. Nowhere is it implied that the color of the sky explains why the weather is the way it is; the only implication is that a particular condition will occur based on a certain observation.

An explanatory model of the weather might involve a series of differential equations deduced from theoretical propositions relating air pressure, relative humidity, wind currents, and the like, and it is quite possible that the predictive success of this model might be less than that of the old sailors' adage. The choice between these two models would depend on one's goals. Looking at the sky might be the best approach if one is interested simply in predicting the immediate weather conditions. If, on the other hand, one wishes to understand the process, then it would be far better to reanalyze the internal logic of the second model in hopes of

refining the hypothesized relationships among components and ultimately producing a higher success rate.

A similar situation exists with models that are used to predict the locations of archaeological sites. If one is simply interested in predicting whether a location will or will not contain a site, then in many areas of the world a highly successful predictive statement would be to say that no individual location contains a site. This conclusion is based on the fact that sites are relatively rare "events" and cover only a minute fraction of the earth's surface. For example, two surveys conducted in conjunction with predictive modeling in the mountainous sections of the western United States showed that in at least these cases a "no site" prediction would have been right more than 99 percent of the time (Kvamme 1983; Reed and Chandler 1984).

Cultural resource managers and archaeologists, however, are less concerned with the overall predictive success rate of a model than with the likelihood of a wrong prediction. Basically there are two types of predictive errors: a prediction can be made that a location (or area) contains a site when in fact it does not, and conversely a prediction can be made that a location does not contain a site when in fact it does. The first type of error may lead to increased costs or to inefficient use of resources and will be called a *wasteful error*. Errors of the second type lead to the destruction of cultural resources and will be termed *gross errors*.

The errors defined above can be associated with the classical Type I and Type II errors defined by Jerzy Neyman and Egon Pearson in a series of papers in the late 1920s and early 1930s (e.g., 1933a, 1933b). As these statisticians pointed out more than a half century ago, in a hypothesis-testing framework there are always two potential errors: we may reject the null hypothesis when it is in fact true (Type I) or we may accept the null hypothesis when it is false (Type II). To relate these errors to predictive modeling, we can take as the null hypothesis that an area will not contain a site. If we reject the null hypothesis when it is true (i.e., accept the fact that there is a site when there is none) we are committing a Type I error or, as it may be viewed from a management perspective, a wasteful error. If, on the other hand, we accept the null hypothesis that a site does not exist in the area when indeed one does, then we are committing a Type II error, which we have more forcefully named a gross error.

An ideal predictive model minimizes both types of errors; that is, it makes correct predictions. In practice, however, models do make wrong predictions. In this regard, we can make two observations. First, in general it is much more costly in cultural resource management to make a gross error than a wasteful one. Second, the likelihood of making a gross error is inversely related to the likelihood of making a wasteful error. To see the logic of the second point, one needs to understand that the primary means of reducing gross errors is by increasing the amount of land predicted to contain sites. But unless site location can be predicted with no errors (which is highly unlikely), this procedure will increase the number of wasteful errors.

The choice between two models, then, has less to do with overall success than with minimizing errors, especially gross errors. In general, a more powerful predictive model is one that for a specific proportion of gross errors to total predictions also minimized the area predicted to contain cultural resources. Let us assume, for example, that there are two predictive models of site location for the same region, Model A and Model B. When both models predict that 5 percent of the region will contain sites, predictions derived from Model A are found to be correct 70 percent of the time, while those from Model B are correct 80 percent of the time. Our first inclination would be to conclude that Model B is a superior predictor. Let us say that upon closer examination, however, we find that in all its predictions Model A makes only 5 percent gross errors while Model B makes 10 percent. For most management purposes, then, Model A is twice as good as Model B (for additional discussion of these two types of modeling errors, see Chapter 8).

TYPES OF MODELS

Until now the discussion has proceeded as though differences in types of models were not important. While it may be true that any model that satisfactorily minimizes errors can be a useful predictor, the form of the model will determine in large part the confidence placed in it and one's willingness to make it even better.

The scientific literature is replete with discussions of models, modeling, and prediction (e.g., Braithwaite 1960; Hempel 1965; Kaplan 1964; Salmon 1971; Scriven 1959, 1962; Zetterberg 1963). During the past two decades archaeologists have also become increasingly interested in these subjects (Binford 1972, ed. 1977; Clarke 1968, ed. 1972; Earle and Christenson 1980; Flannery 1968, ed. 1976; Fritz and Plog 1970; Gardin 1980; Read 1974; Renfrew 1973; Renfrew and Cooke 1979; Renfrew et al. 1982; Salmon 1975, 1976, 1978). Archaeological models range from simple analogs to complex simulations. Although the properties and forms of the various types of models differ in important respects, a more fundamental distinction, which bears directly on any discussion of the types of models used to predict site location, can be made.

In general, models can be divided into two groups based on the degree to which they can be operationalized. Those that contain components or relationships between components that cannot be measured in a replicable and reliable manner will be termed *intuitive* models, whereas those with components that can be so measured will be called *objective* models. Objective models are further distinguished on the basis of *(a)* the spatial referent of the dependent variable (i.e., whether aspects of site location for an area or specific locale are being predicted), *(b)* the predominant form of procedural logic (inductive or deductive), and *(c)* the nature of internal relationships among model components (i.e., whether independent variables are given equal weight or relative weights). On the basis of these criteria, three categories of objective models can be defined: *associational, areal,* and *point-specific* models (Table 3.1).

TABLE 3.1.

Types of objective predictive models of site locations

Model Type	*Primary Procedural Logic*		*Variable Weight*	*Spatial Referent*
	Inductive	*Deductive*		
Associational	Overlay or composite models	Adaptive types	E Q U A L	A R E A L
Areal	Map interpolation Pattern recognition Grid prediction	Simulation Discrete probability distributions Hierarchical decision models	R E L A T I V E	
Point-specific	Pattern recognition Point-specific prediction	Central place models Gravity models Optimum location models Polythetic-satisficer models		P O I N T

The classification presented above differs from the one presented in Chapter 2. The previous typology was based on three criteria: the level of measurement of the independent variables, the model's procedural logic, and the target context. Here our primary concern is not with the level of measurement but simply whether the measurements are made in a consistent and replicable way. For models that can be operationalized in an objective manner, interest now shifts to the form of the model, that is, to the relationships among the internal components and to the nature of the dependent variable.

Intuitive Models

Intuitive models can be derived through either inductive or deductive logic, with the reference frame being either the archaeological record or patterns of human behavior. An example of an intuitive model is the statement, "You'll find arrowheads on high ground near water." This statement may be based on repeated observation or on a common-sense theory about human behavior. But regardless of whether the statement is based on inductive observation or on deductive thinking, the most important characteristic of this model from a scientific standpoint is that the components are not fully conceptualized. While everyone may understand the thrust of the statement, there will not necessarily be agreement on what is high ground or what "near water" means. The relationship(s) among landform, distance to water, and artifacts is only partially established. Until everyone can agree on what the terms mean they cannot be operationalized in a way that is replicable.

Until the variables are operationalized they cannot be measured, and without measurement the relationship(s) cannot be tested.

Many archaeologists might contend that intuitive models are not really models at all, reserving that term only for constructs that can be measured and tested. Leaving aside the philosophical issues, there is good reason to consider intuitive thought in a discussion of predictive modeling. Much of the recorded archaeological data base in the United States was derived through intuitive models. American archaeologists have only recently concerned themselves with formalizing their notions about site location into research designs. Many archaeologists have surveyed and continue to survey land based on their ideas about where they will find sites. Moreover, these intuitive models are often the basis for more intensive research projects. For example, in the early 1970s the Corps of Engineers began plans for the development of Sardis Lake, a reservoir covering about 1400 ha in southeast Oklahoma. The agency commissioned a survey that consisted of one person trying to find as many sites as possible in a 1-month period (Neal 1972). The survey was based on personal intuition and reports from amateurs and resulted in 31 sites being recorded. These sites, along with six others recorded later, formed the basis for 10 years of intensive excavation.

Not only have intuitive models been the basis of much professional work, they have been the mainstay of amateur archaeology. As a result, recorded site locations in most areas of the United States do not necessarily reflect where sites are located but only where people have looked for them. Models of site location based on existing data can lead to predictions with very high accuracy rates. After all, if people have only looked for sites in certain types of places, then it is inevitable that site locations will be highly correlated with specific environmental attributes. This is not to say, however, that all data collected on the basis of intuition must be ignored. Procedures for reducing the biases inherent in this type of data do exist (e.g., subsampling and weighted analysis) and will be discussed in Chapter 7.

It is important to remember that intuitive models are not examples of bad science or of bad thinking. Indeed, creativity and intuition are the most important and most illusive parts of the scientific process. The first question many archaeologists ask themselves prior to designing a survey for a region is, "If I were a prehistoric inhabitant, where would I live?" The problem is that many archaeologists stop there and never formalize their answer. Thus, no matter how brilliant their insight or how many sites they find, no one can objectively evaluate how well their model works.

Objective Models

Associational Models

Often archaeologists are interested in determining whether patterns exist in the data. For instance, suppose a survey records 25 sites in a 1000 ha piñon-juniper

zone and 10 sites in an adjacent 2500 ha sagebrush zone. The first question asked by an archaeologist might be whether the difference in site frequency between the vegetative zones is greater than would be expected if there were no association between site location and the vegetation. One approach to answering this question would be to compute a goodness-of-fit statistic. If the value obtained exceeded a specific level of a chi-square distribution, the association could be considered significant.

If it were determined that a significant association existed, the results might be used as the basis for a simple predictive model. It might be predicted, for example, that in another study area more sites would be found in the piñon-juniper zone than in the sagebrush zone. If this prediction were based solely on the patterning observed in a single survey, our confidence in it would be fairly low regardless of the strength of the association or the proximity of the two study areas. Confidence in the expected outcome might be greater if this prediction were based on 15 surveys in nearby regions, although we still would not be in a position to express our confidence in a quantitative fashion.

Models similar to the one described above are common throughout archaeology. Many predictive models developed in cultural resource management studies take the form of relatively simple pattern-recognition, associational models. For instance, Kohler et al. (1980) conducted an intensive survey of the Halloca Creek drainage, which consists of about 2 percent of the area of the Fort Benning Military Reservation in Georgia. Twenty-one prehistoric and 10 historical sites were found. Site locations were examined to determine whether they covaried with six environmental variables. To evaluate the relationship between soil type and site location, for example, the observed numbers of sites per soil type were compared with the distribution expected if there was no relationship. After computing the appropriate chi-square statistic, the investigators concluded that the relationship between site distribution and soil type was nonrandom.

In a similar fashion Kohler and his colleagues examined the associations between site location and vegetation, distance to water, slope, relative elevation, and distance to roads. The results suggested that the distribution of sites was nonrandom in relation to slope, soils, and horizontal distance to water and that it was random relative to the other variables. For each significant environmental feature, the investigators defined a variable with two states, favorable to site location and unfavorable to site location. A map of each variable was created for the entire military reservation, along with a composite map on which the three variables were overlaid. Areas where favorable values for all three variables intersected were considered high-probability zones; areas with two favorable scores were defined as medium-probability zones; and the remaining areas were considered low-probability zones.

Associational models like the one described above are among the most commonly used predictive models in cultural resource management (e.g., Campbell et al. 1981, 1983; Chandler et al. 1980; Grady 1980; Klesert 1982, 1983; Larralde and Nickens 1980; Reed and Nickens 1980; Thomas et al. 1981). These models are

attractive primarily because of their simplicity; they are easy to construct and relatively straightforward to understand. They are not without their problems, however. For one thing, it is simply not true that the intersection of several favorable values for environmental variables will necessarily be a better predictor of archaeological site location than the individual variables themselves. The intersection is only a more precise predictor if the variables are independent of one another, which is highly unlikely with environmental variables. For instance, well-drained soils are only associated with certain types of landforms and with a restricted number of vegetative communities. Each of these variables *individually* may be highly correlated with site distribution, but before it can be concluded that the predictive power of the model will be increased by using all three simultaneously it has to be shown that site distribution is associated with each variable after controlling for the influence of the other two (see Chapter 5 for an extended discussion of spatial autocorrelation and statistical independence).

A second major problem with associational models is generalization. For the most part, associational models have been developed as part of Class I overviews or using the results from surveys of management-selected areas. They are usually not derived from probabilistic sample surveys and thus may contain biases that will be magnified if the model is generalized (i.e., extended to areas that have not been surveyed).

The predictive power of this type of model, and certainly the generalizability of associational models, would be increased if the suggestions concerning the associations between site location and environmental attributes were not based solely on pattern recognition but instead were deduced from principles of human behavior. One would then be in a position of demonstrating that an association between site location and an independent variable or set of independent variables exists, as well as being able to explain why the association exists.

From a research perspective, explanation is our ultimate goal; only when we can explain why the phenomenon occurs can we be said to truly advance our understanding of human behavior. Deductively derived models, however, are also superior from a management point of view. If we do not understand why patterns occur, our confidence that they will reoccur in the future will always be somewhat tempered. This is especially true when we deal with human behavior. The assumption that settlement locations were conditioned by environmental features may be valid in a general sense, but it will not explain why sites are frequent in one river valley and rare in another. Pattern-recognition models often show that settlement distributions are highly patterned, but without some sort of explanatory framework, management decisions based on these patterns are grounded more on faith than on reason.

There are only a few examples of deductively based associational models. One such model was developed by Sabo et al. (1982; see also Sabo and Waddell 1983) in a cultural resources overview for the Ozark–St. Francis national forests in Arkansas. These investigators used the concept of *adaptation type* to model successive prehistoric and historical human ecosystems in the Ozark Mountains. An adaptation type

relates regional environmental potential to specific levels of socioeconomic and technological organization. Sabo et al. (1982) defined four prehistoric and 14 historical types in the Ozarks, e.g., Late Pleistocene/Early Holocene hunting and gathering, and Late Holocene horticultural, hunting, and gathering adaptations. Expected archaeological site types and their distributions within four major physiographic zones were derived for each adaptation type. The predictions were tested with 254 previously recorded sites. For each site, attributes of four environmental variables were recorded. Q-mode cluster analysis resulted in groups that corresponded to the predicted site classes.

In general, the Ozark–St. Francis model is more convincing than a pattern-recognition associational model, but it would be even more convincing if the adaptation types were not so broad. One cannot avoid the sinking suspicion that, given the conceptual framework, virtually any result could be viewed as consistent with the model. The general approach, however, is in the right direction.

Has the emphasis on associational models in cultural resource management contexts really been misplaced? The answer seems to hinge on the stated objectives. Associational models provide a means of operationalizing the environmental variables that may be related to site location. In this sense they are a tremendous improvement over intuitive models. Associational models can be used to provide a first guess about site location and as a basis for future research; they can, for instance, define environmental dimensions that will be useful in stratifying a region for a Class II survey. Associational models, then, can be a good first step, but hardly a step at which to stop.

Areal Models

Areal models are those that predict certain characteristics of sites or cultural resources, such as density or frequency, per a specified unit of land. For the most part, areal models are more attractive than associational models because the latter only produce relative statements about site location, such as "more sites will be found in this area than in that one" or "more sites are found in this zone than would be expected by chance alone," and these statements are often inadequate for research or management needs. In many instances researchers and managers want to know more than just the fact that one zone will contain more sites than another; they want to know how many sites each zone will contain and what the site density in each zone will be.

Answers to such questions lie in the area of *estimation,* that is, deriving a reasonable estimate of an unknown characteristic of sites and/or of site distribution in a specified region on the basis of a sample of that region. This issue falls under the topic of sampling, which will be discussed in more detail in Chapter 6. Because of the close association between sampling and many forms of areal models, some archaeologists have viewed predictive modeling as synonymous with sampling for the purpose of parameter estimation (e.g., Ambler 1984). There are, however, good reasons for keeping the two separate. Parameter estimates are based on assumptions

about how the population of characteristics is distributed and how that population is sampled. When some type of probabilistic sampling design is used (i.e., when the sampling technique, frame, fraction, unit, etc., are specified), an estimate of a population value can be computed. While this value is the best guess or prediction of the population value, it must be remembered that it is not characteristics of the populations that are being modeled but characteristics of the sampling distribution. That is, sampling theory makes no statements about how the population was derived (in this case, about how sites become located in specific places). Instead, sampling theory only allows us to determine the likelihood that a particular sample would result given a certain hypothesis about the underlying population.

Predictive models of site location (as they are being defined here) all use some aspect of site location as the dependent variable that is being predicted by one or more independent variables. In areal models the nature of the relationship between the independent variable(s) and the dependent one is usually determined for relatively small areas, and this same relationship is then projected to exist in larger, more inclusive areas. Although this notion of projecting from a sample to a larger population is similar to parameter estimation, many areal models are generalized on some basis other than probability theory.

Kriging, for instance, is a technique for generalizing that uses the concept of spatial autocorrelation—the presence of a characteristic in one area makes its presence in adjoining areas more likely (see Chapters 5 and 7). Basically a method of map interpolation, kriging uses moving averages and involves estimating values, and the errors associated with those values, for spatially distributed variables. Although kriging has been most extensively used in trend analysis on geologic mineral deposits, Zubrow and Harbaugh (1978) have provided several examples of how this technique can be used to predict site densities on the basis of samples. In one example they simulate how an archaeologist can divide an area into grid units and then, using his/her intuition about where sites are located, survey 12.5 percent of the grid units most likely to contain sites. A krige analysis of the results produces site density estimates for the entire region that are reasonably close to the true values.

Kriging and other map interpolation techniques, such as trend surface analysis, have been largely ignored as bases for predictive models in cultural resource management, probably because most archaeologists are not well versed in these techniques. Whatever the reason, it is fair to say that the potential of models based on map generalization has not been realized. Models of this type could be especially useful at the Class I or overview stage of work (e.g., Hansen 1984).

One of the most popular types of predictive models used in cultural resource management is an areal-based pattern-recognition model. Although differing in form, most of these models utilize sample data to compute a mathematical function, which is then used to predict some aspect of site location (e.g., presence/absence or site density) for unsurveyed units. A variety of statistical techniques have been used in these models, including multiple linear regression, discriminant function

analysis, and logistic regression, but whatever the statistical technique the logic of these models is the same.

The predictive model developed for the Bisti–Star Lake region of northwestern New Mexico (Kemrer 1982) is a good example of this type of model. The Bisti–Star Lake region is located in the San Juan Basin and consists of various tracts of coal leases totaling approximately 191,500 ha (77,500 acres). The modeling approach adopted was to use the results of six previous surveys to create a predictive model. On the basis of these results, Kemrer and his associates devised eight site classes (Table 3.2), each of which served as the dependent variable in a separate predictive model.

Landsat multispectral satellite data were then used to classify soil and water-source characteristics of the area into eight "environmental classes." Adopting an approach similar to that often used in remote sensing, the investigators used a sample of training pixels (in this case equivalent to an area of about 50 by 70 m) with known environmental characteristics to obtain a mathematical function by which unknown pixels throughout the area could be classified. In this manner very fine scaled environmental data were obtained.

The next step was to place a 2 by 2 km grid over a map of the Bisti–Star Lake region. Seventy-eight "environmental" variables were then calculated for each grid square. Eight of these were simply the number of pixels per unit for each of the eight environmental classes. A second set of eight variables consisted of the proportion of pixels per grid square classified into each class. The remaining 56 variables represented all unique two-way interactions between frequency and proportional variables, respectively, of the eight environmental classes.

TABLE 3.2.

Bisti-Star Lake region site classes

Site Class	*Description*	*Number*
Lithic	undiagnostic lithic scatters	410
Anasazi	sites dated to Basketmaker III - Pueblo III, as well as all sites considered to be Anasazi but not assigned to phases	178
Pre-1933 Navajo	Navajo sites dating from the late 1600s to 1933	146
Post-1933 Navajo	Navajo sites dating from 1933-1980	358
Total Navajo	all Navajo sites combined (includes those that could not be assigned a date)	569
Anglo/Spanish	historical sites dating from 1700-1940	3
Unknown historical	historical sites that could not be affiliated with a specific group	14
Total	all sites combined	1174

From Kemrer 1982:62

The 2 by 2 km grid squares were then used as units of observation for which the dependent variables (the number of archaeological sites of a specific site class per unit) and the independent variables (the 78 variables based on different methods of associating pixels of each spectral class per unit) could be measured. Linear equations were developed for each site class using a multiple linear regression formula. In essence, these equations served as predictive models so that if the values for the eight environmentally related pixel variables could be determined for a grid unit, the number of sites of each site class could be predicted.

The models were tested with data derived from a 15 percent sample survey of the Bisti–Star Lake region. Areas surveyed were not chosen through a probabilistic sampling design but were instead purposely selected to test the entire range of variability in cultural resource density. Based on the discrepancies between predicted and observed numbers of sites, the models were refined by recalculating the linear equations with the survey data.

Models such as the one described above have recently become very popular in cultural resource management (e.g., Gordon et al. 1982; Kranzush 1983; Lafferty et al. 1981; Morenon 1983; Nance et al. 1983; Newkirk and Roper 1982; Peebles 1983; Sessions 1979). Much of this popularity is probably due to the ease with which these models are created and to their apparent predictive power. Two inherent problems of these models should be pointed out, however. First, as with many spatial analytic techniques, grid size affects the results. The models developed on the basis of a 2 km grid in the Bisti–Star Lake region differed substantially from those based on 1 km squares in nearby regions (compare Kemrer 1982 with Sessions 1979). Studies in other areas have also shown that widely differing results can be expected as the grid size is altered (e.g., Kranzush 1983), and thus far no one has been able to resolve this issue for a particular region, to say nothing of the general case.

A second problem, which is also related to grid size, has to do with the characterization of the environment. Most often the environment of each unit is characterized on the basis of one or, at the most, a small number of points in each grid unit from which environmental variables are measured. These points are argued to be representative of the environment of the larger grid unit. This approach is difficult to justify even for small units (say 40 acres or less) and simply misleading for large units. Commonly this approach leads to inaccurate predictions. For instance, Kranzush (1983) found a relatively high frequency of sites in 40-acre units that were predicted not to contain any. She notes that in many cases the center point of the unit (from which the environmental variables were extrapolated) may not have been suitable for settlement but one could usually find at least one spot in the unit that was suitable.

The approach developed by Kemrer (1982) for the Bisti–Star Lake region is an innovative solution to this problem (see also Tipps 1984), but the use of Landsat images to create environmental variables is not without its difficulties. The development of an environmental data base at pixel-level resolution requires not only appropriate aerial photographs but also a detailed understanding of the statistical procedures involved. For instance, Landsat classes that can be accurately mapped

are often extremely broad owing to the poor spectral resolution of the sensor. Areal models based on such classes, then, may be of little use to the land manager. Yet, even when fine spatial resolution is achieved the information is often wasted because the pixel data have to be aggregated into larger units so that they can be comparable with other independent environmental variables mapped at cruder resolution. These and other considerations will be discussed at length in Chapter 9.

Although inductive procedures, such as map interpolation and pattern recognition, are the most common bases for areal models, such models can also be developed deductively on the basis of theoretical propositions about human settlement. Hierarchical decision models (Limp 1983a, 1983b), simulation models (Chadwick 1978; Thomas 1972, 1973), and probability distribution models (Hodder 1976; Hodder and Orton 1976; Thomas 1972, 1973) are all examples of areal models of this type. As a group these models are more diverse than other categories previously discussed. Although they vary widely in their internal logic and procedure, they do share a common emphasis on explaining why humans settle in certain areas and not in others.

Theory-based, deductive areal models have not received much attention in cultural resource management studies, probably for one or more of three reasons. First, theoretically based models require more time to create. The internal connections between variables must be explicitly stated, as must the logical arguments supporting those relationships. Second, validation procedures are more onerous. Deductive models must demonstrate that they are not only consistent with the data but also more parsimonious than any alternative. In contrast, inductive models are judged primarily on the accuracy of their predictions. No claim is necessarily forwarded about how the population was formed in these models, only that the dependent variable covaries with one or more independent variables.

Some archaeologists contend that all pattern-recognition models are based on the assumption that the environment shapes decisions about where humans settle; this assumption is almost always implicit in these models, and the relationship between environment and human settlement is never specified. Although theory-based models also assume a relationship between environmental factors and settlement, the relationships between environmental factors and locational behavior are spelled out according to some behavioral theory. Thus, these models are easier to critique than those based on the generalization that environment is related in some unspecified way to settlement.

Finally, the predictive statements derived from some types of deductive models are not of a form that is useful for management purposes. For instance, probability distribution models yield statements about the expected number of sites per sample unit, but this type of model will not predict which units will contain sites. Instead, such models predict that in the aggregate a specified number of units will contain no sites, a certain number will contain one site, and so on.

While the three reasons cited above may account for the less extensive use of deductively based areal models in cultural resource management, they are not good

reasons. The fact that deductively based models need to be more explicit and are more difficult to validate should not necessarily be viewed as a detriment. Clearly, a model that has successfully gone through this process has much more research utility than one that has not. Even from management's perspective, there is good reason to keep a balance between inductive and deductive modeling. Inductive models, as they are currently used in cultural resource management, may provide useful day-to-day information. They are not, however, designed to provide deep insight into the relationship between humans and the environment. Yet it is these latter relationships that underlie, albeit implicitly, all inductive models being used. In contrast, deductive models have not to date performed well in providing on-the-ground information for making management decisions. But research into these models is one of the prime mechanisms of forwarding our understanding of man-man and man-land relationships that affect the spatial arrangement of human settlement. Emphasizing one approach to the exclusion of the other is the surest way to stymie the potential of predictive modeling in general.

Point-Specific Models

In the past few years there has been a growing trend to shift the level of prediction from the sampling unit to the site itself. Instead of making predictions about the number of sites in a sampling unit, archaeologists have explored methods of assessing the likelihood that any particular spot will or will not contain a site. The appeal of such an approach to both management and research is immense. Not surprisingly, point-specific models have become the predominant form of site locational modeling within the BLM's cultural resource management program (e.g., Burgess et al. 1980; Kvamme 1983; Larralde and Chandler 1981; Peebles, ed. 1981; Reed and Chandler 1984).

Pattern-recognition point-specific models in archaeology are based on procedures developed in the field of remote sensing (see Chapter 9 for an in-depth discussion of this subject). In remote sensing, scientists use reflected radiation values to classify locations of interest on the earth's surface into prespecified groups, such as forest vs nonforest, wheatfields vs nonwheatfields, and so on. In the simplest terms, they first calibrate a "training set" of known cases, such as vegetation types, by measuring different spectral bands; then for other cases, locations with unknown vegetation types, the different spectral characteristics are used to infer vegetation types. The validity of such classification schemes is evaluated using test data that were not used to calibrate the original model.

A similar approach has been adopted in archaeology, using numerical classification techniques like discriminant function analysis and logistic regression. The predetermined groups are defined on the basis of certain combinations of discriminating variables, so that if the same variables are measured for an unknown case it can be placed with a specified degree of probability into one group or another.

In addition to adopting the numerical classification techniques, many archaeologists have also borrowed the concept of a binary response variable. That

is, given locations are classified as being a *site* or a *nonsite*. This is unfortunate, because all sites are lumped into one category. No distinction is made between big and small sites, functionally distinct sites, or sites from different time periods. This is unrealistic, since nearly all anthropological studies indicate that a particular configuration of environmental variables is not equally important in all temporal and functional contexts. With the site/nonsite dichotomy, however, sites are either present or absent, and all sites are created equal. From a managerial perspective, information on different types of sites may not only be important but required. Clearly, different management strategies are required for small lithic scatters and for large ceremonial centers.

There are also statistical problems associated with lumping all sites into one group. These will be discussed at length in Chapters 5 and 7. Suffice it to say here that these problems fall into two groups. The first has to do with the use of what are usually heterogeneous groups in mathematical models that assume that the groups being used are internally homogeneous. For example, discriminant analysis is a popular modeling technique in which two or more groups are statistically distinguished from one another. If there is only slightly more between-group variation than within-group variation, the results will be largely useless and can even be highly misleading. Lumping site classes together almost always increases within-group variability of the site group, often to such a degree that sites are more dissimilar to each other than they are to nonsites.

The second set of problems involves generalization. In the case of areal pattern-recognition models using probabilistically selected sampling units, generalizing the results is relatively straightforward. The sampling unit is the same as the sample element, and parameter estimates can be computed following formulas for element sampling. This is not the case for point-specific models, since the sites found within the sampling units are used as the units of analysis. Thus, the sample is a cluster sample, and unless the appropriate adjustments are made in calculating the group variances and covariances, there are likely to be serious errors in the computation of the mathematical function (see Chapter 6).

The preceding discussion does not mean that *all* pattern-recognition point-specific models are inaccurate or lead to invalid predictions. Given the strong appeal of these models and the recent emphasis placed on them, however, it is important to discuss the problems that can arise. One solution to some of these problems would be the development of a response variable with multiple categories. The creation of multiple groups does not invalidate the use of such techniques as discriminant analysis or logistic regression. It simply makes them more realistic, flexible, and amenable to management and research concerns. The problem of generalization can be mitigated by careful attention to how the model will be used. If its sole purpose is to act as a heuristic device, pointing out patterns of covariation between the environment and site location, then problems associated with generalizing the results are probably not as critical as if the predictions were to be used as the basis for management decisions.

From a theoretical standpoint, the most powerful locational models should be those that not only predict where sites are located but explain why they are located there as well. Models of this type include central place models (Berry 1967; Berry and Pred 1965; Christaller 1966, 1972; Crumley 1976; Haggett 1965; Johnson 1977; Losch 1954; Skinner 1977; Smith 1976), gravity models (Crumley 1979; Haggett 1965; Johnson 1977; Olsson 1970; Plog 1976), optimal location models (Wood 1978), and polythetic-satisficer models (Williams et al. 1973). Some of these, such as central place or optimal location models, have a long history in the field of human geography and have only recently been adapted for use with nonindustrialized societies by archaeologists and anthropologists. Others, such as the polythetic-satisficer model, have been developed by archaeologists on the basis of ethnographic research and basic principles of human behavior.

Much like deductively based associational and areal models, deductive point-predictive models have been overwhelmingly ignored in cultural resource studies. Many of the reasons for this situation cited in the previous sections also hold true at the point-specific level. These models are more difficult to develop than correlational models, and the validation process is more involved. In addition, the accuracy of these models is usually not very high. For instance, in archaeology central place models are generally used more as a yardstick to evaluate deviations from a theoretical pattern than as a predictor of actual site location.

The land manager reading this section may well have decided that, given the inherent difficulties associated with the use of deductive models, the current emphasis on pattern recognition represents a conscious decision on the part of archaeologists. This is a false impression. Outside the confines of cultural resource management, pattern-recognition models have been much less discussed or developed than their theoretically based counterparts. The reason for this disparity goes beyond any simple explanation of academic vs nonacademic research goals. What appears to have happened is that a perception has developed among landholding agencies that locations of archaeological sites can be predicted within acceptable accuracy levels. This perception was probably fostered by a number of theoretical studies, sponsored at least in part by the BLM and the Forest Service, that investigated the potential of pattern-recognition approaches to predicting site location (e.g., Cordell and Green 1983; Grady 1980; Hurlbett 1977; Kvamme 1983).

The net result has been a tremendous emphasis on the methodological issues involved in prediction at the expense of studies of behavioral processes. The implications of this trend can be illustrated with a simple example. Let us suppose that on the basis of environmental attributes 70 percent of all site locations in a region could *potentially* be predicted. Let us further suppose that an associational model was developed that predicted 50 percent of the site locations. By creating an areal-based discriminant function model the result might be to increase the model's predictive capability to 60 percent; with a point-specific logistic regression model, to 65 percent; and with a point-specific quadratic discriminant model, to 67 percent. The point is that the increase in the sophistication of the statistical models has not led to a proportional increase in our ability to predict site locations. In this case, as

with most of the research in predictive modeling in the past few years, all of the effort has been devoted to finding ways of increasing our predictive power through statistical methods. It is not surprising, then, that as more and more research has gone into predictive modeling this research has yielded smaller and smaller increases in predictive power. Because patterns of environmental attributes will only account for so much of the variation in settlement patterns, no matter how much time and money are invested in developing statistical methods or sampling designs, at some level a point of diminishing returns is reached. That point is rapidly approaching in predictive modeling.

A legitimate question for a land manager to ask would be, "Is the additional 30 percent worth it?" There is no simple answer to this question, although an example from anthropology may be useful. In a study of political systems in highland Burma, Edmund Leach (1954) began with an analysis of the ecological situation. He argued that the distribution of two economic systems covaried fairly well with differences in environmental settings, but that once the environmental correlates had been factored out, a number of differences between systems were still left unexplained. Leach used his ecological analysis as a springboard into a more detailed study of the social structure. The result was a far-reaching (and now classic) analysis of political and social dynamics embedded in a culture, a result that simply could not have been obtained through the study of ecological relationships alone.

As Chapter 4 will make clear, much of what is considered important about the study of archaeological remains is part and parcel of the percentage for which pattern-recognition models cannot account. Although these models are useful and informative in certain contexts, it is also true that no matter how the term is defined, much of what archaeologists consider to be "significant" begins where pattern recognition leaves off.

THE MODEL-BUILDING PROCESS

It should be clear from the foregoing discussion that there are many kinds of predictive models of site location. Some are largely or wholly operationalized, others are intuitive; some are based on deductive arguments, others are inductive. Numerous modeling techniques exist, and the choice of a technique depends on research objectives and the available data base. Moreover, predictive models are not mutually exclusive. As archaeologists have learned over the past decade, the line between induction and deduction is neither hard nor fast. There is no reason why different modeling techniques cannot be used to analyze the same data, and in fact, there is good reason to do just this.

Regardless of the form of a model or of the specific techniques used, the basic steps in the modeling process are the same for all models (Figure 3.1). The rest of this chapter will be devoted to outlining this process; Chapters 6–8 will discuss this process in much greater detail.

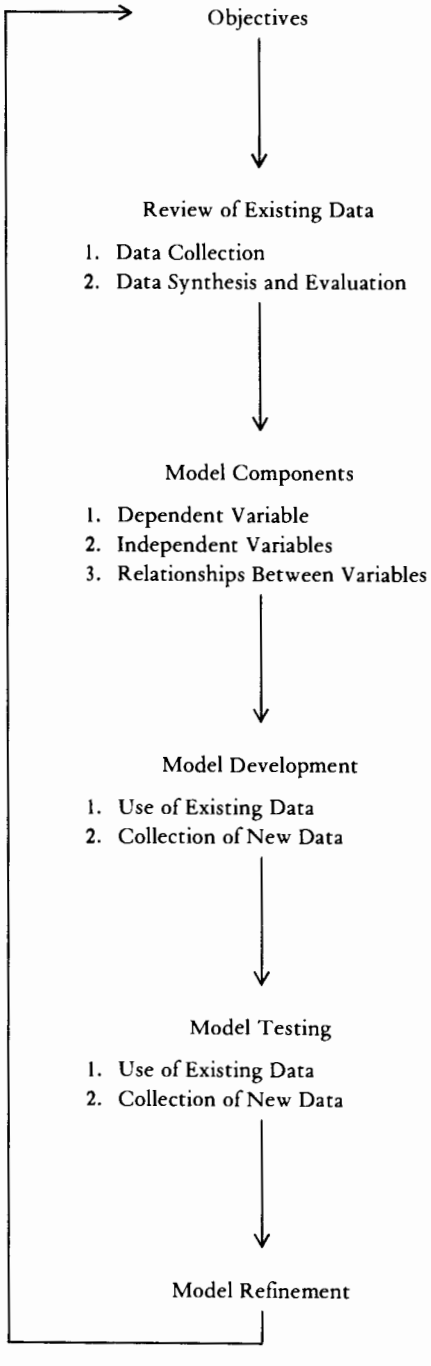


Figure 3.1. The model-building process.

Identification of Objectives

As Figure 3.1 shows, the first step in the modeling process is the specification of goals or objectives. In the process of identifying objectives a clear distinction needs to be maintained between short-term and long-term goals. In the long run, management and research goals are probably not that different; cultural resources are protected for what they can tell us about the past and how the past evolved into the present. It is the information content of the resources, not their physical make-up, that has been deemed worthy of preservation. To best fulfill this legal obligation, federal and state agencies need to know not only where resources are located but also why they are located there. This objective is our end goal. It is not at all clear that we can ever reach it, but as scientists we are committed to continue striving for it.

In order to reach this goal, we need to have a better understanding of the necessary intermediate steps or short-term goals. Often developers of Class I and Class II models refer to their results as "preliminary predictive models," which suggests that they view these models as intermediate steps along the way to a better understanding of site location. Perhaps the most significant criticism that can be made about predictive modeling programs in most cultural resource management contexts is that there is no consensus as to the overall objective of these programs. Models continue to be developed as if they represented the desired end product. Instead of calling for the refinements of existing models, scopes of work usually require the creation of a new model. The results are not cumulative, and thus it is little wonder that most federally sponsored predictive modeling programs are bogged down in a seemingly endless progression of virtually identical models.

From the perspective of the land management agencies, it would be prudent to identify both long-term goals and the steps needed to achieve them. On the basis of this overall plan, an agency could decide whether it would be more productive to award a contract for an overview that requires the creation of a multivariate model of site location or whether it would be more useful to invest that effort in research designed to develop locational variables that make sense from a theoretical standpoint.

Data Collection

The first step in modeling locational behavior for a specific region is to amass the available data. Four basic sources of data are commonly used: historical documents, ethnographic research, archaeological data, and environmental data.

Historical documents include explorers' and colonial accounts of Native American culture and associated settlement patterns. Land-use records are sometimes available, as are baptism and death records for Spanish missions. The latter are especially useful for examining such issues as intergroup movement, population change, and ethnohistoric settlement patterns (e.g., Munoz 1982). Many of these

records have been examined by ethnohistorians, and secondary sources exist for nearly every region of the United States.

Ethnographic research represents a complementary data source. Ethnographic analogy of one form or another has been a mainstay of archaeological interpretation since the inception of the discipline. Ethnographic analyses of indigenous subsistence and settlement systems were used by archaeologists as the basis for settlement-pattern studies long before cultural resource predictive modeling became an issue. Perhaps the best known study of this type is Julian Steward's (1938) *Basin-Plateau Aboriginal Sociopolitical Groups*, which served as the foundation for numerous settlement and subsistence models both within and outside the Great Basin (Flannery and Coe 1968; Jennings 1957; MacNeish 1964; Thomas 1972, 1973; Williams et al. 1973). In addition to direct analogy, ethnographic studies are useful as sources for general propositions about settlement decision behavior (e.g., Jochim 1976; Lee and DeVore 1968; Yellen 1977). Finally, the growing field of ethnoarchaeology continues to supply much-needed data on factors and constraints leading to decisions about where people live as well as on depositional and post-depositional processes that affect the archaeological record (Ascher 1962; Binford 1976, 1978a, 1978b, 1979, 1980, 1981; Coles 1973; Gould 1978, 1980; Kramer 1979).

Recorded archaeological data exist in a variety of forms. Site records are stored at the state level, either in a central repository or dispersed among several state institutions (usually museums and universities). Several federal agencies keep their own records, which may or may not be duplicated at the state repository. Regional data bases, such as the Southwestern Anthropological Research Group (SARG; Euler and Gumerman 1978) and Intermountain Antiquities Computer System (IMAC; University of Utah et al. 1982), exist for some areas. Private institutions, museums, and local historical and archaeological societies also may have information. Finally, as has been true since the beginning of archaeological research, one of the best sources for site locational information is the local informant.

Extant archaeological data vary considerably in quality and quantity. In order to assess the existing data one must evaluate a number of factors. The number and intensity of surveys has a direct bearing on the distribution of known sites and the types of sites recorded. Definitional criteria for sites are often subjective and nonreplicable. The reliability and comparability of recorded information is an open question that must be resolved before this information can be used in model building (see Chapter 7).

Environmental data can be gathered at two levels. At a macro or regional level, data can be collected on a variety of topics, including climate, vegetation, geology, hydrology, and physiography. Sources of these types of data include many federal and state agencies, such as the Soil Conservation Service, the Forest Service, the U.S. Geological Survey, the Fish and Wildlife Service, the National Oceanic and Atmospheric Administration, and the Bureau of Land Management. An increasingly important source of data on environmental conditions is aerial imagery. Remote sensing and Landsat images have emerged as extremely useful tools for

identifying and classifying environmental dimensions and as means of objectively measuring environmental variables (see Chapter 9).

At the site level we are often interested in which environmental features affected the decision to settle in a particular spot. Studies of this nature are classified under the rubric of catchment analysis (e.g., Higgs and Vita-Finzi 1972; Jarman et al. 1972; Roper 1979; Vita-Finzi 1969, 1978; Vita-Finzi and Higgs 1970). Environmental zones that surround each site can be analyzed in terms of their potential economic value. Studies of this type, coupled with environmental and subsistence data from excavated sites (e.g., pollen, flora, fauna, and malacological analyses), can help to shape our understanding of the subsistence strategy.

Data Synthesis and Evaluation

Once the available data have been gathered, they must be synthesized and evaluated in terms of their applicability for predicting site location. One of the first tasks is to identify general trends of cultural change and stability and trends in the distribution of known sites. Map interpolation techniques, such as trend surface analysis, kriging, etc., can often be useful aids in discerning general trends.

One result of this type of background research must be the identification of known sites or at least of the types of sites crucial to understanding regional settlement systems. Here interest lies in determining the effects of what some authors call the "big site" phenomenon (Rogge and Lincoln 1984) and what will be called "magnet" sites in Chapter 6. Implied in the notion of a magnet site is the existence of social factors that led people to locate other types of sites closer to or farther from a particular site than would be expected just on the basis of the prevailing subsistence system. Unless the exact locations of these magnet sites are known, it is extremely doubtful that site locations can be successfully predicted in that region.

In the Santa Cruz River Valley of southern Arizona, for example, a predictive model was developed on the basis of a Class I overview (Westfall 1979). A Class II sample survey demonstrated that the Class I model overestimated the importance of certain environmental zones and therefore was not particularly useful. A second predictive model, which was based on environmental variables derived from work in the Gila Bend area about 80 km (50 mi) to the west (McCarthy 1982), was also tested against the Class II results and again was found not to be a very accurate predictor of site location. An intensive Class III survey revealed the problem; three major Hohokam communities were identified in environmental contexts that did not contain such communities in the Gila Bend area. Each community consisted of a central platform-mound complex surrounded by smaller sites lying within 1.5–5 km of the central complex (Rogge and Lincoln 1984). Only a small proportion of sites were found outside these communities.

In most areas of the country the proportion of known, large, complex sites is higher than the corresponding proportion of known sites in other categories. People

have been drawn to large sites, especially those that exhibit major architectural features or mounds, since the nineteenth century. Many of these sites, which probably represent social centers and/or the top elements of the regional settlement hierarchies, have been formally recorded or are at least known to local residents. The point is that in areas where socially complex societies developed, predictive models based solely on environmental variables are bound to fail. Yet, in most areas the locations of many of the magnet sites are known and can be determined either by examining the existing site records or asking local informants. Thus, the existence and importance of these sites can be evaluated at an early stage in the modeling process (say a Class I level). If this were accomplished, the construction of useful social predictive variables should be possible.

This discussion of magnet sites points out the importance of being able to distinguish site classes. Ideally, site classes are defined along two dimensions, time and function. In practice, however, this task is often difficult even with excavation-based data, to say nothing of the problems involved in using site files or even survey-based data. At the data-evaluation stage it is important to determine (or hypothesize) the types of sites expected to be found for each culture period and their probable locations. The magnitude of the discrepancy between theory and existing data can then be gauged. That is, we can determine how many sites can be classified by period and function, with the remaining sites grouped into a residual category. Examination of the residual category, which in many areas of the western United States will constitute between 60 and 80 percent of all recorded sites, will determine the types of research questions that can legitimately be asked. These questions in turn will affect the type of dependent locational variables that can be modeled and thus the nature of the independent variables that can be used.

Identification of environmental dimensions along which site locations vary is an important step. It is, however, only one step. Most predictive models developed in cultural resource management contexts have viewed this step as the *only* one or at least the most important one, paying lip service to other factors affecting site location. It is also important to bear in mind that the environmental variables that directly covary with site location are probably best viewed as proxies for whatever decision-making criteria led to the selection of locations exhibiting this environmental feature (Kohler and Parker 1986). For example, landform may be a proxy for considerations of defense, agricultural potential, floral resources, or any other reason that a group may have for choosing a place to live or to conduct activities. It follows that several environmental variables may reflect the same decision-making criterion or that one environmental variable may be an indicator of portions of two or more decision criteria. Moreover, the criteria for choosing a site location were probably different in different parts of a single settlement system, and certainly these criteria changed through time and between settlement systems.

Given this situation, it would be best to study the covariation of environmental features with each separate site class. This ideal situation is rarely realized because of the problems of distinguishing site classes, but it is still possible to model expected distributions of sites based on theoretical principles or ethnographic cases

and evaluate the results against the known data base. Using ethnographic information about Great Basin settlement systems, Thomas (1972, 1973) wrote a computer simulation that projected the expected distribution of cultural remains across environmental zones and then tested these predictions against the archaeological record. This approach offers a way of evaluating the effects of environmental attributes on the settlement system that could be a powerful complement to the pattern-recognition studies in vogue today.

In addition to examining environmental factors that affect decisions about where to settle, we need to evaluate the natural processes that affect the creation and present state of the archaeological record. Archaeologists have become increasingly sensitive to the difference between the systemic context in which residues of past behavior are deposited into the archaeological record and the archaeological context in which they are recovered (Ammerman et al. 1978; Binford 1976, 1978b, 1979, 1980; Ebert et al. 1984; Schiffer 1968, 1976; Schiffer and Rathje 1973; see also Chapter 4 of this volume). In general, this growing awareness has not been incorporated into predictive models, probably because of our poor understanding of these processes and of the attendant difficulties in modeling them. Failure to take into account depositional and postdepositional processes leads to predictive models that, at best, predict where sites have been seen and not necessarily where they are or were.

Several recent studies indicate the potential for increasing the power of predictive models by including geomorphic factors. For example, Artz and Reid (1983) use a relatively simple soil-geomorphic model to predict the location of buried Archaic sites in the Little Caney River Basin of northeastern Oklahoma. Previous surface surveys had not found any Archaic materials in the area, leading some investigators to question whether the region had been occupied during this period. Artz and Reid developed a model based on the proposition that the relative age and stability of a geomorphic surface is often reflected by the properties of the soil developed below it. The model was used to identify buried surfaces that in the past were suitable for habitation. Subsequent investigation of these surfaces showed that Archaic sites, although buried, were indeed located in the Little Caney River Basin.

Model Components—Dependent Variables

To develop a model one has to be clear about exactly what it is that is being modeled. As far as site location is concerned there are a variety of potential dependent variables. It is possible to predict site presence or absence, site density, site types, site functions, or various combinations thereof. Moreover, the dependent variable can change, although this will require either drastic internal revisions or an entirely new model. For example, at an early stage of research archaeologists might predict that sites will be found in greater numbers in areas within 100 m of

permanent water and on land with slopes with less than a 5 percent grade. Formally this relationship might be expressed as

$$P(A|B \cap C) > P(A)$$

where $P(A)$ stands for the probability that an area contains a site, B for areas within 100 m of permanent water, and C for land with slopes of less than 5 percent grade. Thus, the equation simply states that the probability that an area contains a site is greater if it meets conditions B and C than it is for all areas in general.

At a later stage of research it may be found that the relationship between site location and the two independent variables is much more precise. This relationship might be modeled with a linear equation of the form

$$A = D + F_1B + F_2C + E$$

where A equals site density; B is distance to water in meters; C is slope in degrees; F_1 and F_2 are the weights for B and C, respectively; D is a constant; and E is an error term. In this case two independent variables are being used to predict the number of sites per survey unit. While the two equations represent two fundamentally different models, it is also fair to say that they are part of the same model-building process, with the latter equation being a more refined expression of the former.

Ideally the dependent variable should be specified first, followed by creation of the model. Usually in predictive modeling, however, a dependent variable is selected on the basis of the data available and the types of independent variables being used. Most archaeologists tend to be less concerned with the exact nature of the dependent variable (as long as it bears on some aspect of site location) than with meeting the assumptions of the modeling procedure, especially in a mathematical model.

In general, we want to proceed from crude measures of site location, such as relative densities (i.e., more sites here than there), to more powerful variables that will predict a specific site type in a particular location. Although the level of locational specificity modeled is directly related to the nature of the data that can be used to test it, it is necessary to guard against blind acceptance of a dependent variable simply because a particular modeling technique is used. Deciding to predict site density because "that's what multiple linear regression predicts" is definitely putting the cart before the horse. Selection of an appropriate dependent variable has to do with defining management and/or research objectives as well as identifying the nature of the data base available or being collected. Once this decision has been made, an appropriate way to model the phenomenon can be found.

Model Components—Independent Variables

Selecting independent variables and determining their interrelationships are perhaps the most difficult steps in the model-building process. There are no rules

that govern this process and few guidelines that can be offered. Variables and their relationships can be derived from inspiration, intuition, creative thought, and/or previous experience. Certainly it is true that if one has a good grasp of general anthropological or sociological propositions about the factors that affect decisions as to where behaviors will be conducted, one is more likely to make an informed choice of variables. There is no guarantee, however, and Clark's (1982:232–234) discussion of false starts and mental gestation periods aptly describes this process.

The development of model components and the definition of their interrelationships should be the areas in which archaeologists make their greatest contribution to the predictive modeling process. This, however, has not been the case. Instead, there has been a tendency among archaeologists producing predictive models to concentrate on the sophisticated multivariate mathematical techniques and to give only casual attention to the predictive variables. In most cases, methodological discussions focus on the inner workings of the statistical procedures with only passing references to the reasons why specific variables were chosen or to how these variables are theoretically related to site location. Indeed it appears that investigators are assuming that the relationship(s) between the environment and site location cannot be specified, other than that there is one, and that if only enough environmental variables are put into the equations something useful will come out.

There is nothing wrong with searching for patterns, but it is important to realize that the ways in which aspects of the environment are conceptualized and measured seriously affect the types of statistical tests that can be used as well as how they are interpreted. Since most archaeologists are more atuned to the relationship between site locations and the surrounding environment than they are to statistical theory, it stands to reason that it is in this area of specification of locational/environmental relationships that archaeologists could make important in-roads.

In an ideal setting a predictive model would be built by first identifying the characteristic of site location, such as site density or frequency (i.e., the dependent variable) and then identifying all the social, environmental, and geomorphic factors (i.e., the independent variables) that impinge upon it. One can envision a series of differential equations describing the relationships among the various factors. In order to learn whether a site would be found at a particular location one would simply assign appropriate values to the variables in the equations, and "presto!" the answer would appear. Unfortunately, at this time such a model cannot be created. While it might be possible to incorporate all three factors into one model, the result would be extremely complex, difficult to evaluate, and probably would have very low predictive power.

Perhaps the best approach for now is to develop a series of models. For instance, it might be hypothesized that settlement in a specific river valley followed some process that can be modeled with a specific probability distribution. The importance of specific environmental variables might be assessed through the use of a pattern-recognition technique. Finally, a model of paleo land surfaces that would have been suitable for habitation could be constructed using information about

geomorphic processes. The results of the models would be mutually reinforcing. If one model worked better than another in a particular area, this information could be used to refine the model and eventually would yield a better understanding of the settlement process.

Regardless of whether one or several models are developed, the form of each model will be the same. In each case a dependent variable will be predicted by one or more independent variables. Some models in archaeology consist of logical statements (such as "if . . . then") that connect the independent variables in some type of causal or deterministic fashion. These models are useful when theoretical reasons can be posited for the connections. Often, however, archaeologists cannot be this specific, and in these cases there are two advantages to using a mathematical model: the relationships between the variables are explicit, and the variables must be objectively defined and measured, a feature often lacking in the logical models.

The major disadvantage of mathematical models is that each model comes with its own set of underlying assumptions. For instance, most of the statistical techniques used in predictive modeling assume a linear relationship between the variables. Theoretically, there is no reason to believe that the relationship between site location and the environment is linear any more than it is quadratic or any other function. While the goal is to work toward theoretically defined connections between variables, a start must be made somewhere, and it is perfectly reasonable to begin this process by using predefined relationships between variables as long as it is understood that these relationships are arbitrary.

Once a specific modeling technique is chosen the necessary data to develop the model must be gathered. For some types of models the data may already be on hand. Associational models can be developed on whatever data exist. The minimal restrictions imposed by these models and the ease with which they can be developed probably account for their popularity in overview-level research.

Other types of models will require the collection of new data or the reformatting of existing data. For example, once it is decided to model site density per kilometer (A) on the basis of slope (B) and distance to water (C) using a linear equation of the form

$$A = D + F_1B + F_2C + E$$

information must be collected on A, B, and C so that the weights (F_1 and F_2), constant (D), and error term (E) can be defined.

The decision as to whether to use existing data or to collect new data to develop the model will depend on the following criteria. Are the temporal and functional site classes that can be defined with existing data sufficient for the model? Are the environmental data that can be obtained from existing maps or site forms suitable for the proposed model? In particular, can patterns in microenvironmental variability be identified from existing records and does the distribution of known sites by environmental zone reflect aboriginal settlement decisions or is it skewed by postdepositional processes? Finally, since many predictive models generalize from a sample, can the existing data be considered in any sense to be representative

of the phenomena of interest? To answer these questions, information must be gathered about the size and distribution of previous surveys as well as their intensities. Using this information, the researcher can determine whether survey results are comparable, if all environmental zones have been adequately covered, and if the types of sites found within the surveyed areas are representative of the settlement system as a whole.

Based on these criteria, gaps in the existing data base can be discerned. In order for the model to be successful, data on paleoenvironmental and geomorphic conditions, chronological and functional dimensions of site classes, and social and economic aspects of the subsistence and settlement systems must meet the requirements of the modeling technique. The existing data base must also be assessed to determine how far the data can be generalized. From this evaluation, the researcher can determine what types of data, if any, must be collected in the field before model building can begin.

Once gaps in the existing data are defined, a research program can be developed to obtain the needed information. While it may seem obvious that research programs should be developed to meet the needs of the particular situation, this has often not been the case. In the usual course of events the first major research project in a region is an overview, combining a review of the existing data and a literature search and producing a planning document (e.g., BLM 1978). In essence, the primary goal of this overview is to decide how future work should be conducted.

It would seem logical that the sample surveys that generally form the next step in these major research projects should be based on the designs outlined in the overview documents. In practice, sample surveys tend to follow rigid, almost standardized formats in which 10 percent of a management-defined area (often an aggregate or series of aggregates of coal lease tracts) is sampled in 40- or 160-acre quadrats through the use of a simple or stratified random sample (see Berry 1984 for a discussion of other problems with this approach).

The uniformity of this design appears to be based on a desire to obtain consistent and comparable results. While the objectives are commendable, the approach is misguided. As will be discussed throughout this volume, the selection of sampling technique, sampling fraction and sample size, and sample unit size and shape are decisions that cannot be made in the abstract but are dependent on the nature of the phenomena of interest and the research objectives. A 40-acre quadrat may be an ideal sampling unit for estimating site density but a very poor choice for studying intersite relationships. Moreover, consistent results have less to do with the sampling design than with issues of survey intensity, site visibility, and sample unit accessibility (see Chapter 6). Indeed, the best approach to achieving substantive comparability between projects is not through design standardization but instead through design flexibility.

The research design not only specifies how the area will be searched for sites but also how sites will be defined and recorded. Definition of site classes will usually require fairly intensive artifact analyses. "No collection" (or limited collection)

policies, while perhaps defensible from a preservation standpoint, run counter to modeling requirements. The present situation in which temporal and/or functional site classes are only poorly developed is unlikely to change unless intensive artifact collections are made. Again, as with sampling design, decisions regarding data recording are best made in relation to a specific project and not at an agency-wide level.

Model Testing

A central aspect of model development is model testing; in fact it can be argued that a model does not really exist until it has been tested. Model testing requires independent data. In general, archaeologists have relied either on collecting new data for testing or on splitting their sample in two, using one half to develop the model and the other half to test it. The former tendency has led to many predictive models remaining untested or being tested only with the data used to derive them. The latter approach often results in such small samples that models can be neither reliably developed nor reliably tested. There are a number of statistical techniques for validating models that circumvent many of the problems described above; these techniques are discussed in Chapter 5.

In the validation stage it is necessary to examine not only the model itself but also the data upon which it is based. Double-blind tests, common in forestry and agriculture, are totally lacking in cultural resource management. Most agencies try to ensure that land is surveyed for cultural resources only once. While the intent of this policy is understandable (after all, if the entire land base can never be completely surveyed, why waste money on resurveying parts of it), it must be remembered that the intended use of a predictive model from the agency's perspective is to allow for useful planning and management decisions about cultural resources in a much larger area. Thus, the argument can be forwarded that, because the model is only as good as the data upon which it is based, time and money spent ensuring the quality of the data are prudent and wise investments.

Model Refinement

Unless 100 percent predictive accuracy is achieved, a model can theoretically always be improved by changing the variables and/or respecifying the relationships among them. It is extremely unlikely that we will ever achieve the high level of predictive accuracy that would imply either complete understanding of past behavior or past behavior that was so deterministically patterned that it can be accurately predicted whether it is understood or not.

The real question for the land managing agency is "how accurate is accurate enough?" The answer to this question depends on the agency and on the research objectives as well as the anticipated results. For instance, a first attempt may explain 60 percent of the variance in site location and indicate major trends in settlement patterning. A researcher might consider this result a tremendous success, while a land manager might view it as a dismal failure.

Much of the above discussion has been phrased in an ideal context, where sufficient time and resources are available. In practice, federal agencies are not in a position to execute grandiose regional survey designs. Instead, federal archaeologists have all they can do to inventory lands to be affected by timber sales and mineral leases. But this does not mean that predictive modeling is some "pie-in-the-sky" scheme dreamed up at the state and regional levels and foisted on district and forest archaeologists. Modeling is part and parcel of what we do as scientists. We cannot evaluate a site located during a timber-sale survey unless it is first placed in some type of scientific context or, if you will, some type of model.

At this time it is less important for the archaeological community to show land-managing agencies how to build accurate models than it is for us to demonstrate the proper use and importance of the modeling process. Scientific models are not like model airplanes; they are not built and then put on the shelf. Yet this is exactly what is being done with predictive models of site location. Archaeologists are being asked to build models that can be used *as is* for the indefinite future.

Scientific models get better as they are refined. Usually as our predictive power increases, our understanding of the phenomena increases as well. Better understanding leads to new and innovative ways of looking at old data and of collecting new data. Often sites are found where it was previously believed there were none, even in areas that have been looked at before.

From a management perspective, the most important issue facing the agencies is not whether to invest in predictive models but whether the modeling process should be an integral part of the overall cultural resource management program. It can be argued that the agencies should utilize models and the modeling process because it is in their best interest to do so. In the short run the first few predictive models will probably not be very powerful. They will not be substitutes for inventory surveys, and perhaps they will not even be very good planning tools. Moreover, a commitment to the model-building process may require the restructuring of the cultural resource management program to ensure that projects are designed to meet specific objectives and that their results are cumulative. Standardization will have to give way to flexibility in research design, and the agencies may have to be prepared for larger rather than smaller sampling fractions. In the long run, however, a commitment to modeling may be the land managing agencies' best hope for the creation of useful tools to guide future development and management of this country's cultural resources. It is to this end that this volume is devoted.

On behalf of Martin Rose and Chris Nagle, my coauthors in Chapters 5 and 6, I would like to acknowledge the following individuals. Throughout the years our ideas on predictive modeling and spatial analysis have been sharpened by conversations with, comments from, and criticisms of many people. Among the most influential have been George Cowgill, Dan Martin, Mike Garratt, Gene Rogge, and Ken Kvamme. With regard to this particular volume we would like to thank those who supplied documents and manuscripts pertaining to predictive modeling. These include Michael B. Schiffer, Bruce Louthan, Chris Kincaid, and Richard Fike. We would also like to acknowledge the support of the various agencies and firms who sponsored much of our research in predictive modeling.

Of note here are Dr. Prentice Thomas (New World Research), Dr. Christian Zier (Centennial Archaeology), Dr. John Hanson (Kaibab National Forest), Gary Stumpf (Bureau of Land Management), Dan Martin (Bureau of Land Management), and Marie Cottrell (U.S. Army Corps of Engineers). Finally, we gratefully acknowledge Lynne Sebastian and June-el Piper for their masterful editorial skills, and Jim Judge for ensuring that this document reached its final form without substantive changes to the authors' ideas or ideals.

REFERENCES CITED

Ambler, J. Richard
1984 The Use and Abuse of Predictive Modeling in Cultural Resource Management. *American Antiquity* 42:140–146.

Ammerman, Albert J., Diane P. Gifford, and Albertus Vorrips
1978 Towards an Evaluation of Sampling Strategies: Simulated Excavations of a Kenyan Pastoralist Site. In *Simulation Studies in Archaeology*, edited by Ian Hodder, pp. 123–132. Cambridge University Press, London.

Artz, Joe Alan, and Kenneth C. Reid
1983 Modeling the Location of Buried Archaic Sites in Northeastern Oklahoma: A Geomorphic Perspective. Paper presented at the 48th Annual Meeting of the Society for American Archaeology, Pittsburgh.

Ascher, R.
1962 Ethnography for Archaeology: A Case from the Seri Indians. *Ethnology* 1:360–369.

Berry, B. J. L.
1967 *Geography of Market Centers and Retail Distribution.* Foundation of Economic Geography Series. Prentice-Hall, Englewood Cliffs, New Jersey.

Berry, B. J. L., and A. Pred, editors
1965 *Central Place Studies: A Bibliography of Theory and Applications.* Bibliographic Series No. 1, with Supplement. Regional Science Research Institute, Philadelphia.

Berry, Michael S.
1984 Sampling and Predictive Modeling on Federal Lands in the West. *American Antiquity* 49:842–853.

Binford, Lewis R.
1972 Contemporary Model Building: Paradigms and the Current State of Paleolithic Research. In *Models in Archaeology*, edited by David L. Clarke, pp. 109–166. Methuen, London.

1976 Forty-seven Trips: A Case Study in the Character of Some Formation Processes of the Archaeological Record. In *Contributions to Anthropology: The Interior Peoples of Northern Alaska*, edited by E. Hall, pp. 299–351. National Museum of Man Series No. 49. Ottawa.

1978a *Nunamiut Ethnoarchaeology.* Academic Press, New York.

1978b Dimensional Analysis of Behavior and Site Structure: Learning from an Eskimo Hunting Stand. *American Antiquity* 43:330–361.

1979 Organization and Formation Processes: Looking at Curated Technologies. *Journal of Anthropological Research* 35:255–273.

1980 Willow Smoke and Dogs' Tails: Hunter-Gatherer Settlement Systems and Archaeological Site Formation. *American Antiquity* 45:4–20.

1981 *Bones: Ancient Men and Modern Myths.* Academic Press, New York.

Binford, Lewis R., editor
1977 *For Theory Building in Archaeology.* Academic Press, New York.

Braithwaite, R. B.
1960 *Scientific Explanation.* Harper Torchbooks, New York. Originally published 1953, Cambridge University Press, London.

Bureau of Land Management (BLM)
1978 *Cultural Resource Management Manual.* Department of the Interior, Washington, D.C.

Burgess, R. J., K. L. Kvamme, P. R. Nickens, A. D. Reed, and G. C. Tucker, Jr.
1980 *Class II Cultural Resource Inventory of the Glenwood Springs Resource Area, Grand Junction District, Colorado.* Nickens and Associates. Submitted to the Bureau of Land Management. Copies available from Bureau of Land Management, Grand Junction District, Colorado.

Campbell, Janice L., Prentice M. Thomas, Jr., and Carol S. Weed
1983 The Right Stuff: Determining Factors in Site Location on Military Reservations. Paper presented at the Southeastern Archaeological Conference, Columbia, South Carolina.

Campbell, Janice L., Carol S. Weed, and Prentice M. Thomas, Jr.
1981 *Archaeological Investigations at the Fort Gordon Military Reservation, Georgia.* Report of Investigation No. 33. New World Research. Submitted to Interagency Archeological Services. Copies available from Interagency Archeological Services, Atlanta.

Chadwick, A. J.
1978 A Computer Simulation of Mycenaean Settlement. In *Simulation Studies in Archaeology,* edited by Ian Hodder, pp. 47-57. Cambridge University Press, London.

Chandler, Susan M., Alan D. Reed, and Paul R. Nickens
1980 *Ecological Variability and Archaeological Site Location in Southwestern Colorado: The Class II Cultural Resource Inventory of the Bureau of Land Management's Sacred Mountain Planning Unit.* Nickens and Associates. Submitted to the Bureau of Land Management. Copies available from the Bureau of Land Management, Montrose District, Colorado.

Christaller, V.
1966 *Central Places in Southern Germany,* translated by C. W. Baskin. Prentice-Hall, Englewood Cliffs, New Jersey.

1972 How I Discovered the Theory of Central Places. In *Man, Space and Environment,* edited by P. W. English and R. C. Mayfield, pp. 601-610. Oxford University Press, London.

Clark, G. A.
1982 Quantifying Archaeological Research. In *Advances in Archaeological Method and Theory,* vol. 5, edited by Michael B. Schiffer, pp. 217-273. Academic Press, New York.

Clarke, David L.
1968 *Analytical Archaeology.* Methuen, London.

Clarke, David L., editor
1972 *Models in Archaeology.* Methuen, London.

Cordell, Linda S., and Dee F. Green, editors
1983 *Theory and Model Building: Refining Survey Strategies for Locating Prehistoric Heritage Resources.* Cultural Resources Document No. 3. USDA Forest Service, Southwest Regional Office, Albuquerque.

Coles, J. M.
1973 *Archaeology by Experiment.* Charles Scribner's Sons, London.

Crumley, Carole L.
1976 Toward a Locational Definition of State Systems of Settlement. *American Antiquity* 78:59-73.

Crumley, Carole L.
1979 Three Locational Models: An Epistemological Assessment for Anthropology and Archaeology. In *Advances in Archaeological Method and Theory*, vol. 4, edited by Michael B. Schiffer, pp. 141-173. Academic Press, New York.

Earle, Timothy K., and Andrew L. Christenson, editors
1980 *Modeling Change in Prehistoric Subsistence Economies.* Academic Press, New York.

Ebert, James I., LuAnn Wandsnider, and Signa Larralde
1984 Theoretical, Methodological and Economic Aspects of Nonsite Surface Survey, Nonsite Sampling and Predictive Modeling. Paper presented at the 49th Annual Meeting of the Society for American Archaeology, Portland.

Euler, Robert C., and George J. Gumerman, editors
1978 *Investigations of the Southwestern Anthropological Research Group: The Proceedings of the 1976 Conference.* Museum of Northern Arizona, Flagstaff.

Flannery, Kent V.
1968 Archaeological Systems Theory and Early Mesoamerica. In *Anthropological Archaeology in the Americas*, edited by Betty J. Meggers, pp. 67-87. Anthropological Society of Washington, Washington, D.C.

Flannery, Kent V., editor
1976 *The Early Mesoamerican Village.* Academic Press, New York.

Flannery, Kent V., and Michael D. Coe
1968 Social and Economic Systems in Formative Mesoamerica. In *New Perspectives in Archaeology*, edited by Sally R. Binford and Lewis R. Binford, pp. 267-283. Aldine, Chicago.

Fritz, J. M., and F. T. Plog
1970 The Nature of Archaeological Explanation. *American Antiquity* 35:405-412.

Gardin, Jean-Claude
1980 *Archaeological Constructs: An Aspect of Theoretical Archaeology.* Cambridge University Press, London.

Gordon, E. Kinzie, Kris J. Kranzush, Laura M. Viola, and Donna J. Knox
1982 *A Class II Cultural Inventory of the Lower White River and Danforth Hills Known Recoverable Coal Resources Areas (KRCRA), Moffat and Rio Blanco Counties, Colorado.* Gordon and Kranzush, Inc. Submitted to the Bureau of Land Management, Contract No. YA-533-CTO-1063. Copies available from the Bureau of Land Management, Craig District Office, Craig, Colorado.

Gould, R. A.
1978 *Explorations in Ethnoarchaeology.* School of American Research and University of New Mexico Press, Albuquerque.

1980 *Living Archaeology.* Cambridge University Press, Cambridge and New York.

Grady, James
1980 *Environmental Factors in Archaeological Site Locations.* Colorado Bureau of Land Management Cultural Resources Series No. 9. Denver.

Haggett, P.
1965 *Locational Analysis in Human Geography.* St. Martin's, New York.

Hansen, Eric
1984 The Whole is Greater than the Sum of its Parts: Small-Scale Surveys and Regional-Level Research. *American Antiquity* 4:35-43.

Hempel, Carl G.
1965 *Aspects of Scientific Explanation.* Free Press, New York.

Higgs, E. S., and C. Vita-Finzi
1972 Prehistoric Economies: A Territorial Approach. In *Papers in Economic Prehistory*, edited by E. S. Higgs, pp. 27–36. Cambridge University Press, London.

Hodder, Ian
1976 Some New Directions in the Spatial Analysis of Archaeological Data at the Regional Level. In *Spatial Archaeology*, edited by David L. Clarke, pp. 223–351. Academic Press, New York.

Hodder, Ian, and Clive Orton
1976 *Spatial Analysis in Archaeology*. Cambridge University Press, London.

Hurlbett, Richard E.
1977 *Environmental Constraint and Settlement Predictability, Northwestern Colorado*. Colorado Bureau of Land Management Cultural Resources Series No. 3. Denver.

Jarman, M. R., C. Vita-Finzi, and E. S. Higgs
1972 Site Catchment Analysis in Archaeology. In *Man, Settlement, and Urbanism*, edited by P. J. Ucko, R. Tringham, and G. W. Dimbleby, pp. 61–66. Schenkman, Cambridge, England.

Jennings, Jesse D.
1957 *Danger Cave*. Society for American Archaeology Memoir 14.

Jochim, Michael A.
1976 *Hunter-Gatherer Subsistence and Settlement: A Predictive Model*. Academic Press, New York.

Johnson, G. A.
1977 Aspects of Regional Analysis in Archaeology. *Annual Reviews in Anthropology* 6:479–508.

Kaplan, Abraham
1964 *The Conduct of Inquiry: Methodology for Behavioral Science*. Chandler, New York.

Kemrer, Meade F., editor
1982 *Archaeological Variability within the Bisti–Star Lake Region, Northwestern New Mexico*. ESCA-Tech Corporation. Submitted to the Bureau of Land Management, Contract No. YA-553-CTO-114. Copies available from Bureau of Land Management, Albuquerque District, Albuquerque, New Mexico.

Klesert, Anthony L.
1982 A Predictive Model of Aboriginal Site Location within the Rio Grande National Forest, Colorado. In *Studies of the Prehistoric and Historic Cultural Resources of the Rio Grande National Forest*, edited by Anthony L. Klesert. Centuries Research, Inc. Submitted to the USDA Forest Service. Copies available from USDA Forest Service, Southwest Regional Office, Albuquerque.

1983 A Predictive Model of Site Potential in the Rio Grande National Forest, Colorado. Paper presented at the 48th Annual Meeting of the Society for American Archaeology, Pittsburgh.

Kohler, Timothy A., and Sandra C. Parker
1986 Predictive Models for Archeological Resource Location. In *Advances in Archaeological Method and Theory*, vol. 9, edited by Michael B. Schiffer, pp. 397-452. Academic Press, New York.

Kohler, Timothy A., T. P. Desjeans, C. Feiss, and D. F. Thompson
1980 *An Archaeological Survey of Selected Areas of the Fort Benning Military Reservation, Alabama and Georgia*. Remote Sensing Analysts. Submitted to the Heritage Conservation and Recreation Service, Contract No. C-5716(78). Copies available from Interagency Archeological Services, Atlanta.

Kramer, C., editor
1979 *Ethnoarchaeology: Implications of Ethnography for Archaeology*. Columbia University Press, New York.

Kranzush, Kris J.
1983 *The Reanalysis of Class II Cultural Resource Inventory Data from the Lower White River Known Recoverable Coal Resource Area, Rio Blanco County, Colorado: Final Report.* Gordon and Kranzush, Inc. Submitted to the Bureau of Land Management, Contract No. YA-553-CTO-1063. Copies available from the Bureau of Land Management, Craig District, Craig, Colorado.

Kvamme, Kenneth L.
1983 A Manual for Predictive Site Location Models: Examples for the Grand Junction District, Colorado. Draft submitted to the Bureau of Land Management, Grand Junction District, Colorado.

Lafferty, Robert H., III, Jeffrey L. Otinger, Sandra Clements Scholtz, W. Fredrick Limp, Beverly Watkins, and Robert D. Jones
1981 *Settlement Predictions in Sparta: A Locational Analysis and Cultural Resource Assessment in the Uplands of Calhoun County, Arkansas.* Arkansas Archaeological Survey Research Series No. 14. Fayetteville.

Larralde, Signa, and Susan M. Chandler
1981 *Archaeological Inventory in the Seep Ridge Cultural Study Tract, Uintah County, Utah, with a Regional Predictive Model for Site Location.* Bureau of Land Management, Utah Cultural Resource Series No. 5. Salt Lake City.

Larralde, Signa, and Paul R. Nickens
1980 Archaeological Inventory in the Red Wash Cultural Study Tract, Uintah County, Utah. In *Sample Inventories of Oil and Gas Fields in Eastern Utah, 1978-1979,* assembled by Paul R. Nickens. Bureau of Land Management, Utah Cultural Resource Series, No. 5. Salt Lake City.

Leach, Edmund R.
1954 *Political Systems of Highland Burma.* G. Bell and Sons, London.

Lee, Richard B., and Irven DeVore, editors
1968 *Man the Hunter.* Aldine, Chicago.

Limp, W. Fredrick
1983a Location Choice and Settlement Predictions: The Conformity of Method and Theory. Paper presented at the 48th Annual Meeting of the Society for American Archaeology, Pittsburgh.

1983b An Economic Model of Settlement Aggregation and Dispersal. In *Ecological Models in Economic Prehistory,* edited by Gordon Bronitsky, pp. 18-45. Arizona State University Anthropological Research Papers No. 29. Tempe.

Losch, A.
1954 *The Economics of Location,* translated by W. H. Wolgam and W. F. Stolper. Yale University Press, New Haven.

MacNeish, Richard S.
1964 Ancient Mesoamerica Civilization. *Science* 143:531-537.

McCarthy, Carol Heathington
1982 *An Archaeological Sample Survey of the Middle Santa Cruz River Basin, Picacho Reservoir to Tucson, Arizona.* Arizona State Museum Archaeological Series No. 148. Tucson.

Morenon, E. Pierre
1983 Environmental Diversity and Prehistoric Site Location: Rhode Island Case Studies. Paper presented at the 48th Annual Meeting of the Society for American Archaeology, Pittsburgh.

Munoz, Jeanne
1982 A Partial Index to the Mission San Gabriel Baptism, Marriage, and Death Registers. Report submitted to the U.S. Army Corps of Engineers, Contract No. DACW90-81-C-0016. Copies available from the U.S. Army Corps of Engineers, Los Angeles District Office.

Nance, C. Roger, Harry Holstein, and David C. Hurst
1983 Evaluation of Multiple Regression Models Predicting Archaeological Site Distributions at Fort McClellan, Alabama. Paper presented at the 48th Annual Meeting of the Society for American Archaeology, Pittsburgh.

Neal, Larry
1972 *Assessment of the Cultural-Historical Resources of Clayton Lake.* Oklahoma River Basin Survey, University of Oklahoma. Submitted to the U.S. Army Corps of Engineers. Copies available from the U.S. Army Corps of Engineers, Tulsa District, Tulsa, Oklahoma.

Newkirk, Judith A., and Donna C. Roper
1982 *Predictive Modelling in the Piceance Basin, Northwestern Colorado.* Gilbert/Commonwealth. Submitted to the Bureau of Land Management, Contract No. YA-553-CT1-136. Copies available from the Bureau of Land Management, Craig District, Craig, Colorado.

Neyman, J., and E. S. Pearson
1933a On the Problem of the Most Efficient Tests of Statistical Hypotheses. *Philosophical Transactions of the Royal Society* 231:281–337.

1933b The Testing of Statistical Hypotheses in Relation to Probabilities A Priori. *Proceedings of the Cambridge Philosophical Society* 24:492–510.

Olsson, G.
1970 Explanation, Prediction, and Meaning Variance: An Assessment of Distance Interaction Models. *Economic Geography* 46:223–233.

Peebles, Thomas C.
1983 Discriminant Analysis in Site Predictive Models: Problems and Prospects. Paper presented at the 48th Annual Meeting of the Society for American Archaeology, Pittsburgh.

Peebles, Thomas C., editor
1981 *Report on the Cultural Resources in the Former Navajo-Hopi Joint Use Area (FY1978).* Department of Anthropology, Northern Arizona University. Submitted to the Bureau of Indian Affairs. Copies available from Bureau of Indian Affairs, Flagstaff Administrative Office, Flagstaff, Arizona.

Plog, Stephan
1976 Relative Efficiencies of Sampling Techniques for Archeological Surveys. In *The Early Mesoamerican Village*, edited by Kent V. Flannery, pp. 136–160. Academic Press, New York.

Read, Dwight
1974 Some Comments on the Use of Mathematical Models in Anthropology. *American Antiquity* 39:3–15.

Reed, Alan D., and Susan M. Chandler
1984 A Sample-Oriented Cultural Resource Inventory in Carbon, Emery, and Sanpete Counties, Utah. Nickens and Associates. Draft submitted to the Bureau of Land Management, Contract No. YA-553-CT2-1080. Copies available from the Bureau of Land Management.

Reed, Alan D., and Paul R. Nickens
1980 The Cisco Cultural Resources Study: A Sample-Oriented Inventory, East-Central Utah. In *Sample Inventories of Oil and Gas Fields in Eastern Utah 1978–1979,* edited by S. L. Larralde, P. R. Nickens, and A. D. Reed. Bureau of Land Management Cultural Resource Series No. 5. Salt Lake City.

Renfrew, Colin, editor
1973 *The Explanation of Cultural Change: Models in Prehistory.* Duckworth, London.

Renfrew, Colin, and Kenneth L. Cooke, editors
1979 *Transformations: Mathematical Approaches to Culture Change.* Academic Press, New York.

Renfrew, Colin, Michael J. Rowlands, and Barbara Abbott-Segraves, editors
1982 *Theory and Explanation in Archaeology: The Southampton Conference.* Academic Press, New York.

Rogge, A. E., and T. R. Lincoln
1984 Predicting the Distribution of Archaeological Sites: A Case Study from the Central Arizona Project. Paper presented at the 49th Annual Meeting of the Society for American Archaeology, Portland.

Roper, Donna C.
1979 The Method and Theory of Site Catchment Analysis: A Review. In *Advances in Archaeological Method and Theory*, vol. 2, edited by Michael B. Schiffer, pp. 119–140. Academic Press, New York.

Sabo, George, III, and David Waddell
1983 Adaptation Type Models and Site Prediction in the Arkansas Ozarks. Paper presented at the 48th Annual Meeting of the Society for American Archaeology, Pittsburgh.

Sabo, George, III, David Waddell, and John H. House
1982 *A Cultural Resource Overview of the Ozark-St. Francis National Forest.* Arkansas Archeological Survey. Submitted to the USDA Forest Service. Copies available form USDA Forest Service, Russellville, Arkansas.

Salmon, M. H.
1975 Confirmation and Explanation in Archaeology. *American Antiquity* 40:459–470.

1976 "Deductive" versus "Inductive" Archaeology. *American Antiquity* 41:376–380.

1978 What Can Systems Theory Do For Archaeology? *American Antiquity* 43:174–183.

Salmon, Wesley
1971 *Statistical Explanation and Statistical Relevance.* University of Pittsburgh Press, Pittsburgh.

Schiffer, Michael B.
1968 Archaeological Context and Systemic Context. *American Antiquity* 37:156–165.

1976 *Behavioral Archaeology.* Academic Press, New York.

Schiffer, Michael B., and W. L. Rathje
1973 Efficient Exploitation of the Archaeological Record: Penetrating Problems. In *Research and Theory in Current Archaeology*, edited by Charles L. Redman, pp. 169–179. John Wiley, New York.

Scriven, Michael J.
1959 Explanation and Prediction in Evolutionary Theory. *Science* 130:477-482.

1962 Explanation, Prediction, and Laws. In *Minnesota Studies in the Philosophy of Science*, vol. 3, edited by H. Feigl and G. Maxwell, pp. 170–230. University of Minnesota, Minneapolis.

Sessions, Steven E., editor
1979 *The Archaeology of Southwest Gallegos Mesa: The EPCC Survey Project.* Navajo Nation Papers in Anthropology No. 1. Window Rock, Arizona.

Skinner, G.
1977 Cities and the Hierarchy of Local Systems. In *The City in Late Imperial China*, edited by G. Skinner, pp. 275–351. Stanford University Press, Stanford.

Smith, Carol A.
1976 *Regional Analysis*, 2 vols. Academic Press, New York.

Steward, Julian H.
1938 *Basin-Plateau Aboriginal Socio-political Groups.* Bureau of American Ethnology Bulletin No. 120. Washington, D.C.

Thomas, David Hurst
1972 A Computer Simulation of Great Basin Shoshonean Subsistence and Settlement Patterns. In *Models in Archaeology*, edited by David L. Clarke, pp. 671-704. Methuen, London.

1973 An Empirical Test of Steward's Model of Great Basin Settlement Patterns. *American Antiquity* 38:155–176.

Thomas, Prentice M., Jr., Carol S. Weed, L. Janice Campbell, and Jeffrey H. Altschul

1981 *The Central Coal II Project: A Class II Inventory of Selected Portions of Carbon, Emery, and Sevier Counties, Utah.* Report of Investigation No. 25. New World Research. Submitted to the National Park Service, Contract No. CX5000-2-0087. Copies available from the Archeological Service Branch, Division of National Register Programs, National Park Service Southeast Regional Office, Atlanta.

Tipps, Betsy L.

1984 *The Tar Sands Project: Cultural Resource Inventory and Predictive Modelling in Central and Southern Utah.* P-III Associates. Submitted to the Bureau of Land Management, Contract No. YA551-CT3-340038. Copies available from the Bureau of Land Management, Richfield, Utah.

University of Utah, Bureau of Land Management, and Forest Service

1982 Intermountain Antiquities Computer System User's Guide: Instructions and Computer Codes for Use with the IMAC Site Form. Copies available from the University of Utah Archaeological Center, Salt Lake City.

Vita-Finzi, Claude

1969 Early Man and Environment. In *Trends in Geography—An Introductory Survey,* edited by U. Cooke and J. H. Johnson, pp. 102–108. Pergamon, Oxford.

1978 *Archaeological Sites in Their Setting.* Thames and Hudson, London.

Vita-Finzi, Claude, and E. S. Higgs

1970 Prehistoric Economy in the Mount Carmel Area of Palestine: Site Catchment Analysis. *Proceedings of the Prehistoric Society* 36:1–37. Cambridge, England.

Watson, P. J., S. A. LeBlanc, and C. R. Redman

1970 *Explanation in Archaeology: An Explicitly Scientific Approach.* Columbia University Press, New York.

Westfall, Deborah A.

1979 *An Archaeological Overview of the Middle and Lower Santa Cruz Basin: A Class I Cultural Resource Survey for the Central Arizona Project–Tucson Division.* Arizona State Museum Archaeological Series No. 134. Tucson.

Williams, Leonard, David Hurst Thomas, and Robert Bettinger

1973 Notions of Numbers: Great Basin Settlements as Polythetic Sets. In *Research and Theory in Current Archaeology,* edited by Charles L. Redman, pp. 215–237. John Wiley, New York.

Wood, John J.

1971 Fitting Discrete Probability Distributions to Prehistoric Settlement Patterns. In *The Distribution of Prehistoric Population Aggregates,* edited by George J. Gumerman, pp. 63-82. Prescott College Anthropological Reports No. 7. Prescott, Arizona.

1978 Optimal Location in Settlement Space: A Model for Describing Locational Strategies. *American Antiquity* 43:303–309.

Yellen, J. E.

1977 *Archaeological Approaches to the Present.* Academic Press, New York.

Zetterberg, Hans L.

1963 *On Theory and Verification in Sociology.* Reprinted. Bedminister, Totowa, New Jersey. Originally published 1954, Almquist and Wiksell.

Zubrow, Ezra B. W., and John W. Harbaugh

1978 Archaeological Prospecting: Kriging and Simulation. In *Simulation Studies in Archaeology,* edited by Ian Hodder, pp. 109–121. Cambridge University Press, London.

Chapter 4

THE THEORETICAL BASIS OF ARCHAEOLOGICAL PREDICTIVE MODELING AND A CONSIDERATION OF APPROPRIATE DATA-COLLECTION METHODS

James I. Ebert and Timothy A. Kohler

This chapter, intended for both managers and archaeologists, discusses archaeological predictive modeling from the theoretical and methodological standpoint. During discussions between the authors and editors of the volume and Bureau of Land Management and Forest Service archaeologist/managers that took place before the book was written, it was suggested that the material contained herein should be directed toward the cultural resource manager. The implication was that managers would not be interested in the sorts of things that archaeologists often produce. This was to be a practical volume, a guide to how predictive modeling can be done and how it should be used—not a compilation of esoteric anthropological theory. Some of those present seemed to be looking for a guide for the manager/archaeologists on how to do "pragmatic" predictive modeling that would cut research costs; others leaned more toward wanting a document that would question the propriety of using predictive modeling for purposes of assessment or mitigation.

Both groups seemed to feel that locational predictive modeling had already been developed in useful form; the problems from their perspective lay in deciding how or whether modeling should be used. It is our feeling that we do *not* know as much as we should about how to do predictive modeling at present; that it is a worthwhile goal to want to understand the process more thoroughly; and that through the proper combination of rigor and research we can probably learn to do such modeling in the near future. But at this stage in our understanding of the modeling process, it would be premature to attempt to produce a guidebook.

In the two years since the original manuscripts for this volume were written, it has become even more apparent that many archaeologists and cultural resource managers want and *need* a guide to predictive modeling. With accelerating frequency, especially during the past year, we have received calls and letters from colleagues (some of whom are archaeologists and some of whom are not) in the remote sensing and GIS fields who are contracting and experimenting with

archaeologists who want to implement predictive models in their study areas. Those colleagues are invariably armed with third-or fourth-generation xerox copies of early drafts of certain chapters from this volume, chapters that purport to tell just how to do predictive modeling. After wading through the pro's and con's of various regression and sampling methods, they suddenly realize that the "modeling" advocated in those chapters has a surprisingly and perhaps dangerously simplistic foundation beneath all of the mathematical discussions.

"Surely there was more to prehistoric human behavior than this implies," said one colleague, himself a Native American trained as an archaeologist, remote sensing specialist, and geographic information system researcher. "This is what we do to map fox or squirrel habitats: look for water and shelter and food and then draw polygons and isopleths around them. Squirrels don't have canteens but Indians did. Do these archaeologists think they know all about how complex past peoples' seasonal rounds were, why they went where they did?"

The authors of this chapter feel, in fact, that we as archaeologists *do not* know all about the complex systemic behavior that must be the basis of archaeological predictive modeling. The theme of this chapter, then, is that while there may be more than one way to *do* predictive modeling once we know how to do it, as suggested elsewhere in this volume, there is only one way to learn how to do it. For those who contend that we already know how to do predictive modeling ("we do it all the time"), this could be rephrased to read that there is only one way to *prove* that we know how to do it. Developing predictive modeling as a tool to aid both archaeologists and cultural resource managers must proceed from a consideration of just what it is that both of these groups want and need to know about.

While some might feel, superficially, that archaeologists want to "explain" while managers just want to know where and what the resource is, we will illustrate that these goals are inseparable. Both must be approached from a *theoretical* standpoint—starting with the consideration of how we believe systems of human adaptation operated in the past and moving logically in the direction of evaluating how the ways we discover, collect, and analyze our data are compatible with learning what we need to know.

Several reviewers of this chapter have protested that we are presenting "just one theory of predictive modeling" here. We would like to make it clear that the terms *theory* and *theoretical* are used here not in any partitive sense (". . . he has one theory and she has another. . .") but rather to indicate where one begins trying to build the framework of ideas and methods, and the hypothetical links between the two, that will be a prerequisite to being able to do predictive modeling, no matter what one means by that. This chapter, then, is about "The Theoretical Basis of Archaeological Predictive Modeling," as opposed to "The Non-Theoretical Basis of Predictive Modeling." What, one might ask, could be meant by non-theoretical predictive modeling? Again, using *theory* to mean the framework by which ideas are evaluated, a non-theoretical approach would be one that begins with an attempt at the "unbiased" interpretation and derivation of knowledge from data, a direction that we will characterize in this chapter as *empirical* predictive modeling.

Empirical predictive modeling, in its simplest form, consists of using the results of site surveys of an area and matching the locations of sites with certain landform features or other indications of past characteristics of the environment. Once these correspondences are noted, the proposition is set forth that more sites will be found in areas where the greatest proportion of previously found sites was located.

In more complex manifestations, empirical predictive modeling breaks previously found sites into functional or other assumed types, derives complex taxonomies of environmental indicators, sometimes specifies multiple working hypotheses about the relationships between these two sets of variables, and applies sophisticated mathematical models (correlation and other associational analyses) to determine which sets of correlations are strongest. Then the same "prediction" is made—that sites will be distributed in unexamined areas the same way (that is, with respect to the same environmental indicators) that they were in the previously explored area.

In a sense, empirical predictive modeling often works—that is, correspondence between the presence of sites and of gross environmental indicators often exist at some level of statistical confidence. Mathematical confidence tests have nothing to do with explanatory confidence, however; they only test the probability of obtaining specific results by chance, given certain characteristics of the samples from which data are drawn. It will be suggested in this chapter that the "success" of some empirical predictive models has as much to do with the ubiquity of the archaeological record across the landscape, and with natural postdepositional processes, as with the realities of the archaeological record.

This chapter will explore in depth the differences between theoretical and empirical predictive modeling. We begin with general properties of human adaptational systems as a first step in an exploration of the processes that anthropological and ethnoarchaeological research suggests are responsible for the formation of the archaeological record. The complexities of human adaptational systems and their "translation" into the archaeological record may make difficult reading for nonarchaeologists, but they are inescapable. In order to learn to apply empirical predictive modeling to the archaeological record, one must "work back" through these complexities, which may be even more difficult than our approach of "working forward" through them.

It will also be suggested that one way to make this learning task—and future empirical predictive modeling, once we learn how—easier and more economical is to fit our data discovery and measurement methods to the things we want to know about. In other words, we need to make our data-collection methods compatible with our goal of explaining complex, multicomponent human systems. One major difference between present-day attempts at empirical predictive modeling and a theoretical approach is that empirical modeling has inappropriate data biases already built in. The data upon which it is based have been cast in terms of sites with various assumed functions. It will be suggested in this chapter that new methods of data collection, based instead upon our ideas about how the archaeological record is

formed and designed to allow the evaluation of inherent biases, may often be helpful in the development of *any* workable predictive model, whether explanatory or empirical.

As was seen in Chapter 2, researchers have been experimenting with empirical predictive modeling for many years and are continuing to do so today. Most of the locational predictions made in archaeology today are statements of empirical correlation. True prediction of archaeological distributions of materials and of their concomitant behavioral and natural causes is a worthy goal and one that is important and necessary for both the cultural resource manager and the archaeologist. Modeling and prediction are integral parts of the scientific explanatory process, as will be illustrated in this chapter. They form a very real part of what archaeologists must do to link their beliefs about the operation and organization of past systems with the observable remains of the archaeological record, and they constitute the only means by which those beliefs can be tested. Cultural resource managers need to know where archaeological materials are located, where they can be found by archaeologists, and what these materials are in order to preserve or otherwise manage them.

The archaeologist and the manager are united in their attempt to arrive at successful predictive models. There may occasionally be talk of theory vs applications, of the research goals of the archaeological scientist being at odds with the pragmatic objectives and responsibilities of the manager. But research cannot be separated from such applications as attempting to predict the locations of archaeological materials. Research provides information about the basic operation of past human organizational systems; the discard of materials from these systems; the incorporation of archaeological materials into what is discovered and seen as the archaeological record; and the ways in which archaeologists discover, measure, and interpret this record. Without this information there is no hope of understanding the mechanisms that create cultural resources. Prediction is not a rote empirical process: its scope encompasses the entire framework of archaeological inquiry and explanation. Archaeologists and managers are partners in cultural resource management and study.

We conclude our introduction with a discussion of what this chapter is and what it is not. This chapter is different from the rest of the book: it presents ideas about how the world works, about the structure of archaeology and anthropology, about the organization of human systems, about the formation of the archaeological record, and about how archaeologists perceive and use that record. Looking at the task of locational prediction from this perspective tends to highlight the difficulty and intricacy of the task, since it soon becomes apparent that a large number of complex considerations can affect the locations and even the degree of predictability of archaeological materials. These are things that must be explored before we can hope to predict successfully and predict with understanding the locations of cultural resources. Although we present methodological suggestions for overcoming some of these difficulties, we risk being regarded as spoilers to the extent that we cannot at this time offer easy fixes for all of the problems we can foresee in locational prediction.

This chapter is not an overview of how people are currently proposing, or attempting, to do predictive modeling; these topics are discussed in other chapters. Instead of focusing on the modeling process, this chapter discusses some of the things that we need to think about (and some of the ways in which we might think about them) in order to perfect the process of predicting whatever we decide to predict. To begin with, we attempt to define the places that modeling and prediction occupy within the explanatory framework of archaeology—that is, what *are* modeling and prediction? What do we *want* (or what do we *need*) to model and predict? The question of research goals is also addressed—what will we have to learn in order to be able to do these things?

Methodological questions are very important in this discussion. The interpretations that we make concerning the archaeological record are probably influenced as much by how archaeologists deal with their data as by what people actually did in the past. How can we collect the appropriate data? How can we ensure consistency and comparability in data collection, measurement, and analysis within and between surveys and other studies? Should or can every researcher have a unique research problem or orientation, or are there general problems upon which we must concentrate, problems of critical importance to the manager *and* the archaeologist? And finally (and perhaps most important from a management perspective), how can we efficiently collect data and do the other research that is necessary if we are to learn how to predict characteristics of the archaeological record and how to give these characteristics meaning in terms of past behavior?

These and many other topics are explored in this chapter. We begin by discussing the framework of archaeological explanation within which modeling and prediction must take place.

PREDICTION, MODELS, AND THE SCIENTIFIC FRAMEWORK OF ARCHAEOLOGY

The archaeological record is a complex amalgam of patterning in material objects created by the organization of peoples' activities in the past and by the intervening cultural and natural processes that have preserved or rearranged these materials since they were lost or abandoned by their past owners. The archaeological record consists solely of patterns that we can see *today*—that is, it is a contemporary phenomenon. It is important to note that these patterns do not ordinarily record a single moment frozen in time that, given the proper expertise, we should be able to reconstruct. In fact, the archaeological record is not ordinarily the simple result of past episodes of individual behavior, and it is only through a scientific, explanatory archaeological framework that we can give it meaning. Nor is the archaeological record a mirror that reflects past behavior in a dark, warped, and incomplete fashion. This is only the case if what we want to do is to reconstruct in microscopic (and normally impossible) detail an instant view of the past. We would argue that this is not the goal of archaeology. The nature and scale of the archaeo-

logical record is such that we will be more successful in understanding it if we consider it not as the reflection of actions of individuals but rather as the cumulative record of an entire system. These systems are not directly embodied in nor are they equivalent to the materials we find in and on the ground. Linking past organizational systems with the archaeological record can only be accomplished through the explanatory framework of archaeology. The only distortions in this reasoning process will exist in archaeologists' models, not in the archaeological record.

Explanation in Archaeology

Explaining things in archaeology is a two-way street, a progression of theory and method. Theory is the way in which we think about things, particularly about the existence, nature, and direction of cause-and-effect relationships, and method is the way in which we go about dealing with data. These two parts of the explanatory process are inseparable, regardless of what the archaeologist wants to explain. In the chart shown in Figure 4.1, some of the links between theory and method in archaeological explanation are shown. This diagram is intended more as a guide to how we might think about the explanatory process than as an indisputable flow chart of archaeological thought, and many other categories in the progression might be acknowledged. The point is that explanation involves both theory and method. In the diagram, one might proceed in either direction—from ideas about human subsistence, settlement, mobility, and technological organization (that is, the organization of systems) to interpretation of patterning in the archaeological record, or vice versa. Linking the two extremes of this diagram constitutes explanation and requires the modeling of a series of intervening processes. These processes transform the ways that people organized their systems into what we see today as the archaeological record. One class of these processes links static archaeological data with the dynamics of past systems; the study of these has been referred to as formulation of "middle-range theory" (Binford, ed. 1977:6–9). In our diagram, this class comprises discard behavior and depositional and postdepositional processes; in its broadest sense, middle-range theory provides guidelines for generating empirically falsifiable outcomes from general theory. Other factors that further remove the patterning we see in the archaeological record from past systemic organization are those introduced by archaeological methodology itself—the ways in which archaeologists recover, measure, analyze, and interpret the archaeological record.

These things that separate high-range theory from the meaning that we assign to patterned data represent complicating factors in attempts to interpret the archaeological record. Moving from one of these complicating factors to another requires qualitative rather than simply quantitative "translation"—that is, the physical archaeological record left behind after the action of each of these factors is of a very different nature than it was before. In the course of this chapter, each of the components of the explanatory archaeological framework will be discussed. First, however, the place of modeling and prediction—the subject of this volume—in the explanatory process must be addressed.

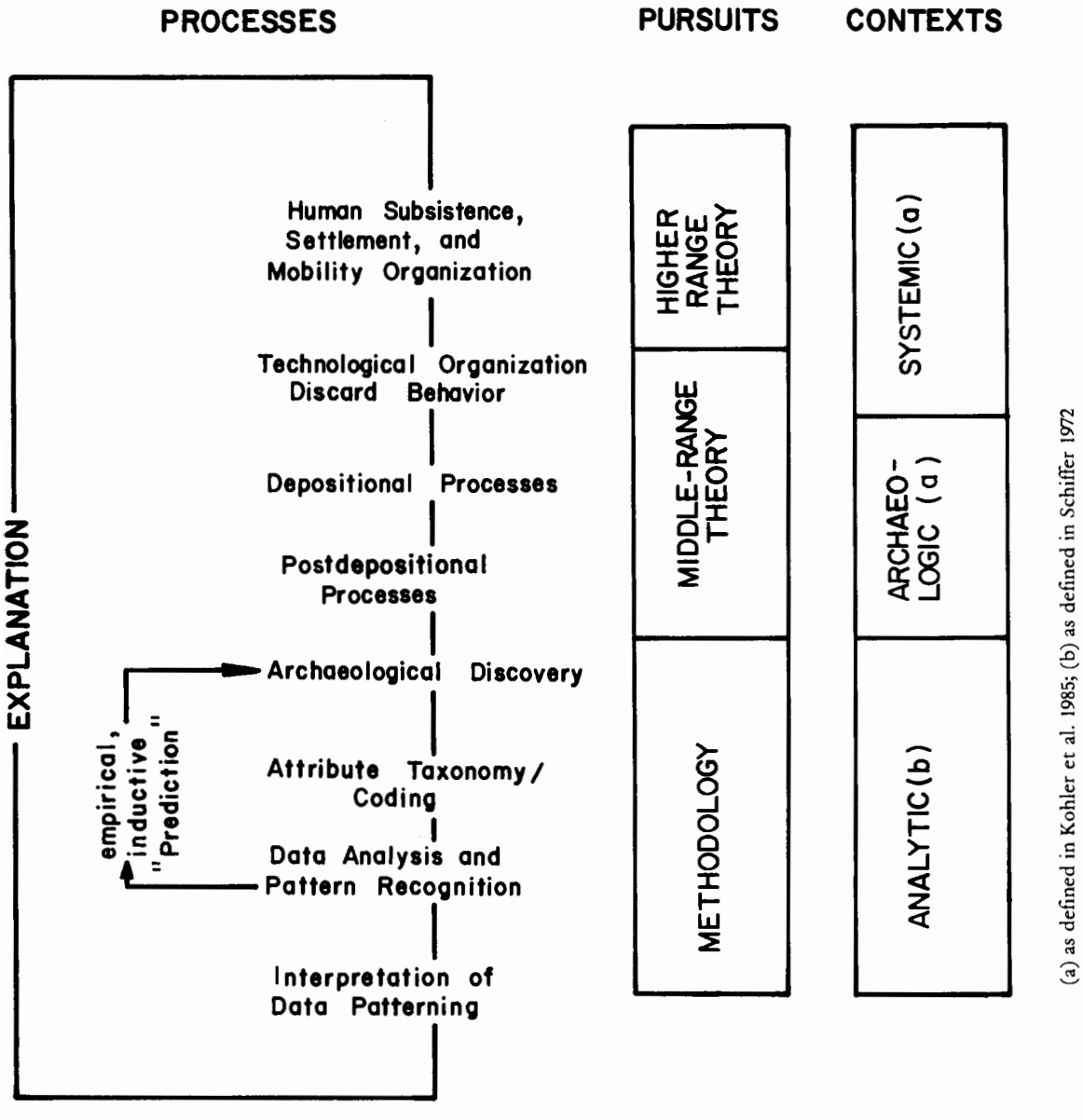


Figure 4.1. The explanatory framework of archaeological science. Explanation is the process of modeling human subsistence, settlement, and mobility organization using archaeological and anthropological data, as well as anthropological, environmental, and systems theory, and confirming these models using prediction to derive expectations for data patterning. These predictions must also be linked with higher-level theory through middle-range theoretical propositions concerning the things that separate the static archaeological record from the organization of human systems. Empirical, inductive projection, sometimes referred to as "prediction" in the literature, is a methodological exercise in which the results of future archaeological discovery are projected from noting correspondences between where sites have been found previously and environmental or landform features of assumed significance.

Modeling and Prediction

In Figure 4.1, the lowest box, interpretation of data patterning, is connected with the highest theoretical category, subsistence and settlement/mobility organization, by the two-way process of explanation. Explanation involves integrating archaeological data with other sorts of information—ethnographic, ethnoarchaeological, historical, environmental—to create *models* that connect the archaeological record with what we think was happening in the past. These models are abstract and complex formulations and can never be proved to be strictly "true." In fact, this is not their purpose: they are constructs that help us to assign meaning, rather than laws or translational rules. Yet they need to be tested or confirmed if we are to know whether they are realistic and useful, and whether they elucidate the mechanisms behind how people live in their world.

The way that models are tested is through prediction. Prediction is the formulation of hypotheses—that is, testable statements of expectations—based upon models. If predictions based on models are found to be successful, then the model and the theories upon which it is based tend to be confirmed. In the structure of scientific explanation, models and theories can never be proved to be true, but if the mechanisms behind the predicted phenomena are being modeled faithfully, the predictions based on them will be consistently successful.

Successful prediction of phenomena in the real world is an accomplished fact in many scientific disciplines, such as electronics, chemistry, and physics. These successes consist of experiments in which predictions based on models are confirmed in a wide variety of situations, with external influences being held "equal." Such successes are unknown at present in archaeology. Not only are we unable to predict phenomena over a wide range of situations, but there is virtually no agreement as to what we want to predict and what we have to model in order to do that.

What Do We Want to Predict and What Do We Need to Model?

The literature dealing with predictive modeling is usually directed toward determining the locations of archaeological materials, whether for discovery purposes (Artz and Reid 1983; Davis 1980a, 1980b; House and Ballenger 1976; Lynch 1980; McManamon 1981a, 1981b; Nance 1980, 1981; Spurling 1980; Warren 1979), for purposes of finding archaeological "voids" (Baker and Sessions 1979; Kemrer 1982; Kemrer, ed. 1982; Klesert 1983; Kvamme 1980, 1982, 1983a; Parker 1985; Peebles 1983; Sabo and Waddell 1983; Scholtz 1980, 1981), or for more avowedly explanatory purposes (Chandler and Nickens 1983; Limp 1983; Nance et al. 1983; Waddell 1983). Prediction of the locations of archaeological materials is a primary concern of cultural resource managers, as well, for in order to manage resources one must know where they are. It could be argued, and will be argued later in this chapter, that prediction of the locations of sites is an ambiguous goal, for the concept of the site is

of uneven usefulness when the ways in which archaeological materials are deposited, accumulated, and discovered are taken into account.

There may well be things other than simple locations, too, that archaeologists and managers might want to predict. Densities of materials, for example, might be of interest (Foley 1981c; Thomas 1973). The diversity or clustering of assemblage components at different sample unit sizes (Whallon 1973, 1974, 1984) or the occurrence of patterning congruent with intrasite activity structure (Kintigh and Ammerman 1982) are other possibilities. The most obvious thing, or at least the first thing, that cultural resource managers need to predict, however, is the location of cultural resources.

To make predictions we need to have models, and those models must span the entire explanatory framework rather than simply concentrating on those things we want to predict. Models exist at a theoretical level, not an empirical one. Their purpose is to elucidate the *mechanisms* behind the formation processes of the archaeological record, i.e., to explain it. Prediction, then, is a subset of explanation. Whether predictions are to be locational or not, it is human organizational systems that must be modeled, as well as all those complicating factors between this highest level of human behavior and the archaeological record as we see and measure it.

Cultural resource managers and archaeologists share the need for explanatory models. We do not yet have many satisfactory archaeological models or even components of such models. It will undoubtedly take many more years to decide what sorts of models are needed by both archaeologists and managers. Some of the things that we may need to consider in this decision process—those "complicating factors" referred to above—are discussed in the remainder of this chapter.

THE NATURE AND ORGANIZATION OF HUMAN SYSTEMS: SETTLEMENT, MOBILITY, AND TECHNOLOGY

A Systems Perspective on Prediction

As discussed in Chapter 2, anthropologists interested in the relationships between people and their environment have increasingly adopted an ecosystemic perspective on these relationships. Over the past two decades archaeologists have also acquired the habit of referring to the dynamic interaction between people and the ecosystem as the settlement system without worrying too much about what it means, in general, to call something a system. (A notable exception is D. L. Clarke [1968].) Yet our acceptance of this term has significant implications for our attempts to predict the locations of cultural resources. A system may be practically defined as

> a circumscribed complex of relatively bounded phenomena, which, within these bounds, retains a relatively stationary pattern of structure in space or of sequential configuration in time in spite of a high degree of variability in the details of distribution and interrelations among its constituent units of lower order [Weiss 1973:40].

This vague characterization can be sharpened by an exclusion. The mere fact that something is composed of several components does not necessarily make it a system; a distinction can be made between *systems* and *mechanisms*. In a mechanism such as a typewriter, for example, one action rigidly triggers other actions in a completely determined manner, corresponding to notions of strict linear cause and effect. In systems, however, there is much freer interplay between the components, despite considerable predictability in the actions of the system as a whole. Systems are not, however, composed of parts that are chaotic in their behavior. Living systems have an evolutionary tendency toward consolidation along stereotyped tracks and toward determinancy in the behavior of the parts; such systems ultimately realize some balance between flexibility (indeterminancy) and rigor (determinancy) (Weiss 1973:54–59). Relatively rigid designs have great efficiency but are only successful if the problems to be solved are always the same.

Another characteristic of living systems (for example, a human community of hunter-gatherers in its regional ecosystem) is that they tend to provide stability of existence for their components (individual bands or households, for example), although the state of any of these components at any time is itself unpredictable, varying far more than the state of the system of which the components are a part (Piaget 1978:59–72). This characteristic of systems leads in turn to a hierarchy of predictability that Weiss calls stratified determinism: there is predictability in the behavior of the system despite demonstrable indeterminism in the individual constituents of that system.

We suggest that human settlement systems share many characteristics with general living systems. Settlement systems are the way that people move around on and locate themselves within a landscape. The individual constituents of this system—the locations of individuals or groups at any given moment, the ways that decisions are made or rationalized, the likes or preferences of human participants, and all the minute details that seem to constitute the everyday world when one is actually *involved* in a system—are inherently less predictable than are the structure and patterning of the system as a whole.

This is not to say that any part of the operation of general systems or of human settlement/mobility systems is random or, in the final analysis, indeterminate. The point is that scientific research addressing a research problem dealing with a system component must be targeted at the system to which the component belongs. Our job in spatial prediction, then, is to understand the structure of the system first. Accordingly, we will spend some time in this chapter discussing what the structure of a settlement system might include.

In the course of this chapter we will argue that modeling undertaken for purposes of predicting the locations and characteristics of phenomena in the archaeological record should take place on the level of human organizational systems. In order to demonstrate this, we propose to take the reader on a journey through the many stages of archaeological explanation, beginning with some approaches to modeling the nature of human settlement/mobility systems.

It will be important to remember that although individuals obviously *make* artifacts and other parts of the archaeological record, neither the patterning nor the role of these portions of the archaeological record in space and time can be equated with the actions (and even less the thoughts or decisions) of individuals or with specific episodes of behavior. Neither are the cultural materials we find today located where they are because of simple interactions between human behavior and specific resources or landscape variables.

The patterning of materials in the archaeological record is a result of the organization of the cultural system that produced those materials. A cultural system is not the summation of the actions of individuals but rather consists of the components in an organizational framework under which actions are structured; the patterning of cultural materials will embody aspects of this framework rather than provide any sort of instant view of a frozen ethnographic moment (Binford 1981).

> In cultural systems, people, things, and places are components in a field that consists of environmental and sociocultural subsystems, and the locus of cultural process is in the dynamic articulations of these subsystems [Binford 1965:205].

The actors in a cultural system are not only people, but places, artifacts, strategies, schedules, landscapes, climate, environment, resources—and many other things as well.

One hallmark of contemporary attempts at archaeological prediction, and indeed of much modern archaeology in general, is the explicit or implicit assumption that environmental factors are major, even exclusive, determinants of much human behavior (site location, subsistence strategies, etc.). Environmental variables, such as distance to water, distance to resources assumed to have been important, shelter, and available lookouts, are compared with the location of archaeological materials to determine whether there are correlations between these landscape characteristics and such cultural variables as the location of sites. The causal link between site locations and natural, independent variables is usually considered to be multivariate—that is, people positioned their sites with respect to an optimal combination of all the resources in which they were interested.

Probably the best example of this approach is in Jochim (1976), often cited as one of the seminal works in archaeological prediction. Jochim argues that, since the distributions of individual resources seldom coincide, these resources exercise differential degrees of "pull" on settlements in relation to their value to the people who occupied those settlements. One problem with this approach is that it is based on a model of the individual person as decision-maker and of specific resources as the basis for making decisions about where to locate activities. That is, it attempts to predict specific components of the larger organizational system without regard to the system of which they are a part. This is not the level of human organization that must be addressed; it is the structure of human organizational systems within ecosystems that needs to be modeled in order to predict things about the components of human systems. How ecosystems variables relate to this task will be

considered later in this chapter; now, however, we will attempt to illustrate an approach to modeling human systems.

Systemic Mobility/Settlement Organization

Archaeologists and cultural resource managers work with an archaeological record produced by prehistoric human systems. All the "facts" that we know about these past systems are actually meanings that we have assigned to the archaeological record. Empirical correlative models that use distances between sites and resources as bases for predictions assume that simple proximity of one thing to another implies some sort of connection or causality, and that distance negates these relationships. The assumption that proximity means something is of course supportable when one is observing ethnographic instants in time. It would be supportable in the interpretation of the archaeological record if we could be assured that we are observing therein instants in past time, i.e., a spatially and temporally nonoverlapping archaeological record. Not only have we no such assurances, but in fact it is almost certain that we are not. Many locations are used for short time periods within most human systems; resources may be transported great distances in anticipation of future needs; and many resources are not in constant demand. Prehistoric people, for instance, could certainly travel some distance without taking a drink, and they certainly had the mental resources to carry water with them. We should have as much capacity to realize (on another level) that the location of one component of a system—where an artifact is discarded, or where a camp is made—is affected by the patterning of other components in space and time: for instance, where another camp was made and what was there last week, or what a group anticipates it will find at the next camp. Rather than being due to the immediate proximity of the resources, in fact, archaeological site patterning is the result of long-term repetition (or lack thereof) in the "positioning of adaptive systems in geographic space" (Binford 1982:6), and the use of space is not uniform, even within the same system. Some activities occur at concentrated locations and some do not. The spatially concentrated nature of some activities and the dispersed nature of others have been discussed in terms of "ranges" of various types (Foley 1977, 1978, 1981b; Jochim 1976), settlements vs activity "nodes" (Isaac 1981:134), and catchments (Vita-Finzi and Higgs 1970).

The very nature of human systems—organized through such tactics as planning and anticipation and effected through caching, transport of materials, staged manufacture, and intensive reuse and recycling of material items—brings the use of proximity arguments in predictive modeling under question. Human behavior is different from animal behavior in that animals in general do not flexibly or consciously anticipate, plan, or transport, cache, and recycle materials; animals do not have behavioral systems organized in a human way.

The things that people do that involve planning, anticipation, and the complex geographic repositioning of materials (some or most of which are not left where

they were used, or are reused there and other places in other times) will not be understandable in any simple way through correlations of artifacts or other cultural evidence and supposed nearby resources. Of course, unplanned events and activities will be represented in the archaeological record, for even in the most highly planned systems (and perhaps particularly in them) unanticipated contingencies will arise. These events, in fact, may be explainable through artifact-resource proximity arguments—these are the things that people do like animals, and the same sorts of predictive modeling that our previously mentioned colleague uses to model fox and squirrel habitats can be used to "predict" them.

So there are aspects of both human organization and "animal" behavior embodied in the archaeological record—perhaps we *do* know how to do some archaeological predictive modeling after all! But before you skip the rest of this chapter and turn to discussions about the best regression models, we think that just a few very important questions must be asked, including *What proportion of human behavior is immediate and unplanned (and thus explainable using proximity arguments) and what proportion is systematically organized? Which portions of human behavior are we most interested in?*

The nature of activities that happen at any place during an occupation will of course have a relationship to the resources available there, but this relationship may not be a simple one, and its strength will be affected by such environmental characteristics as the distribution or diversity of resources (Harpending and Davis 1977) or the annual range of temperatures requiring, enabling, or restricting storage of foodstuffs (Binford 1980). But economic resources are not the only actors in human organizational systems, and they will not be the only determinants of where different activities are carried out in these systems. What a group does at one place, for instance, may be as much affected by what they will do at the next place they visit, or what they did at the last place they visited, as it is by the available resources at the current location.

An examination of one taxonomy of differential mobility patterns will help to illustrate the interlocking nature of the parts of a human organizational system, as well as the implications of different forms of organization for the formation and ultimately the predictability of the archaeological record. Binford (1982) has distinguished a number of ranges or mobility zones that can be used in different combinations to characterize the ways that people use the space around their residential base. A *residential base* is the place where a group lives, where resources are consumed, where children are reared, and where most maintenance activities take place. Residential camp sizes vary, mostly in relation to population sizes. There are certain complications in this relationship, however, that prevent direct projections of population on the basis of site size, as will be discussed later in this chapter. Surrounding the residential base is the *foraging radius,* which is usually considered to be within 10 km of the camp in any direction; resources in this zone are exploited in the course of trips that last a day or less and from which both resources and people return to the residential camp. This area contains *locations,* places where resources are extracted and where limited processing is carried out.

Few maintenance activities are carried out at locations. Outside the foraging radius is a *logistical radius,* which is exploited by special-purpose task groups who stay away from the residential base for at least one night and sometimes for months. Within the logistical radius, both maintenance activities and special-purpose activities can and do take place.

Not all groups use these different radii to the same extent. The use of these radii varies with the frequency with which a group's residential base is moved, and this, in turn, is conditioned by environmental and perhaps in some cases social factors. In highly diverse environments almost all resources can be found within a group's foraging radius, and people in equatorial jungles and possibly in some other environments, such as the Kalahari Desert and the southern parts of the North American Great Basin, particularly during the summer months, acquire nearly all resources using a generalist *encounter* strategy during daily walkabouts. Intensive use of the foraging radius, however, leads to quick depletion of resources, and when this happens the residential camp is moved, most often to one edge of the old foraging radius. From this new basecamp a new foraging radius is established (Figure 4.2). Only half of this new radius is actually usable for foraging, of course, since the portion shared with the old radius is still depleted. This sort of mobility strategy results in what Binford (1982:10) calls a *half-radius continuous pattern.*

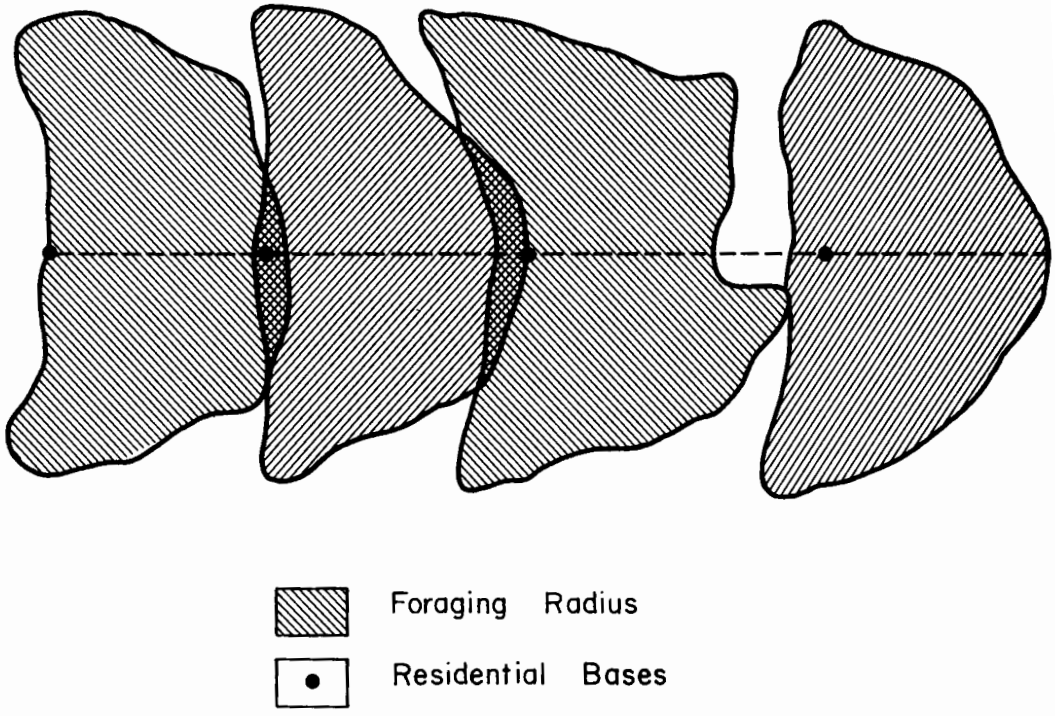


Figure 4.2. The half-radius continuous pattern of exploitation of the landscape by foraging groups. When the resources within a foraging radius are depleted, the group moves its residential base to the far edge of that foraging radius and begins to exploit another half-radius. Materials lost or discarded within the foraging radius are expected to be of low density and relatively continuous distribution (after Binford 1982).

In more differentiated or simpler environments, a *complete radius leapfrog pattern* of residential mobility is more often found ethnographically. This settlement system consists of residential moves that result in little or no overlap between successive foraging radii but produce logistic radii that do overlap (Figure 4.3). In this situation, logistical camps are often located at old residential bases because materials in these abandoned camps can be reused and because the specialized task-group members are familiar with the old residences and their surroundings—reasons for site location that are at least partly nonenvironmental. Examples of cultures with this type of settlement system include the northern Paiute and the Shoshone. A variation of the complete radius leapfrog pattern that is common in lower-biomass settings is the *point-to-point* pattern found in high-elevation settings and claimed to be used, for example, by the Yaghan of Tierra del Fuego (Wills 1980). In this case residential moves involve no overlap in use zones at all, not even in the logistic radii. The location of residential camps under this mobility pattern represents a compromise among the locations of known but spatially incongruent resource distributions. These resources are then exploited through logistic mobility.

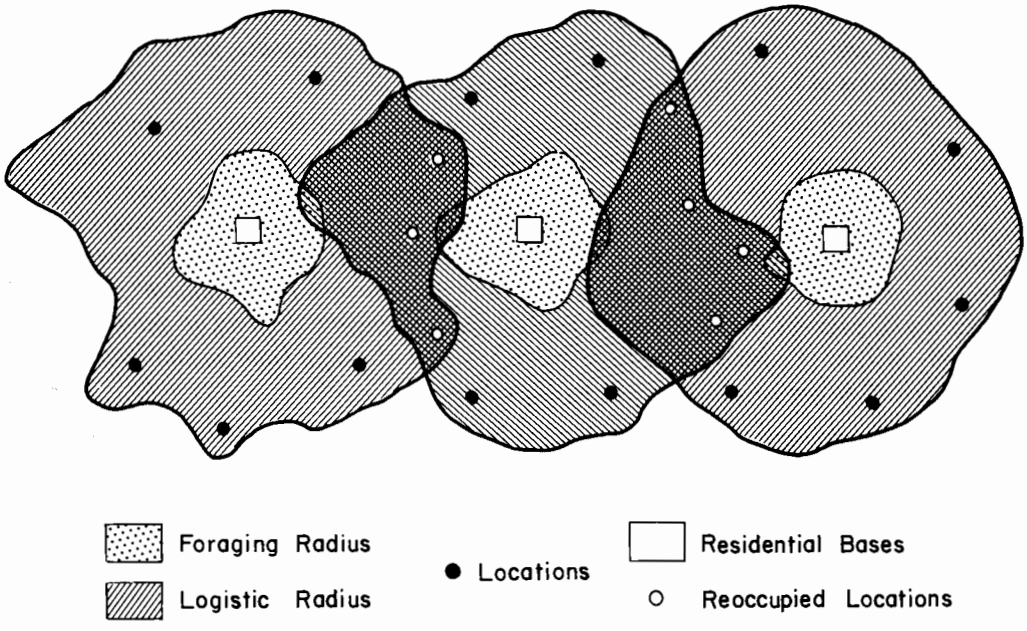


Figure 4.3. The complete radius leapfrog pattern of landscape use. This model was devised to typify the land-use strategy of logistically oriented groups. Locations that are reused within the zones of logistic radius overlap could contain assemblages representing different functional uses. Archaeological materials found within the foraging radii would be dispersed and continuous; materials at locations within the logistic radius are more focused but may represent multiple functional occupations (after Binford 1982).

Some Examples of Variability in Reuse of Places

Binford's mobility/settlement type taxonomy, as described above, is not an attempt to arrive at any "whole truth" about human spatial organization; rather it is an attempt to model different types of organization so that their consequences in the archaeological record can be predicted. Binford's model is not altogether theoretical in its derivation; rather, it is based on ethnographic examples gathered by anthropologists and ethnoarchaeologists studying hunter-gatherers, pastoralists, and agricultural groups throughout the world. Ethnographic and ethnoarchaeological accounts of variations in mobility and settlement patterning indicate that groups operating under different mobility/settlement patterns exhibit different patterns of reuse of places. This observation has important implications for our understanding of the complexity of the archaeological record.

A number of expectations or predictions about the reuse of places can be drawn from Binford's mobility model. Binford's suggestion that under the complete-radius leapfrog pattern old residential bases will be reused for special-purpose logistic functions leads to the expectation that, under such a mobility organization, sites will occur at definite points within the landscape where different functions would overlap. In addition, since the location of residential bases represents a compromise among the locations of resources exploited through logistic mobility, we might also anticipate reuse of residential locations as residences, assuming stable distributions of logistically exploited resources.

For the half-radius foraging pattern, on the other hand, there are no logistic camps and resources are more evenly distributed. Reuse of residential camps might be less common under this form of organization, in part because foraging radii would more likely be depleted of critical resources for some time and in part because of the nature of the environments in which foraging is most commonly practiced, as will be discussed below. Foraging radius locations—places where resources are encountered and perhaps minimally processed—could be expected to occur almost randomly within the foraging radius, a pattern that through time would lead to a low-visibility but continuous archaeological record.

Anthropologists and archaeologists have found that living hunter-gatherer and pastoralist groups that pursue a relatively generalist strategy and fall toward the foraging end of the mobility/settlement scale utilize their foraging radii more or less continuously. Population densities among such groups are characteristically low. An annual average density of 0.03 persons per square kilometer has been recorded among the /Kade area Bushmen (Harako 1978; Tanaka 1969), and even among the relatively densely populated Ituri Forest Pygmy a density of only 0.2 to 0.6 persons per square kilometer is typical. Characteristically, such peoples exploit their sparsely populated ranges relatively evenly.

Foley (1981c:21) cites very low densities of artifacts among such groups in Africa, even on residential bases if those bases were only occupied once. What is more, a large percentage of artifacts among such groups are discarded at what Foley

calls "secondary home range foci," which are the equivalent of Binford's "locations" within the foraging radius. These locations are usually used only once, and their occurrence throughout the environmentally diverse home range assures even distribution of discarded items. Gould (1980) reports that among Australian aborigines only about 1 percent of lithic discard occurs at the residential basecamp; most of the rest occurs within the home range (foraging radius). The results of evenly distributed, low-density discard over the length of time monitored by ethnologists are almost invisible, but over archaeological time this discard process can produce impressive and relatively continuous densities of discarded materials.

This discussion has important implications for the ways in which the archaeological record of foragers should be surveyed, measured, bounded, and analyzed, a topic to be discussed in greater length in later sections of this chapter. Given a foraging adaptation, it is clear that, in much of the contemporary archaeological record, discrete "sites" will not be apparent. Nonetheless, the continuous archaeological record left by groups employing a foraging strategy includes within it materials related to both types of activity areas used by these groups (residential and nonresidential loci).

Although few human groups pursue a pure foraging subsistence strategy, most groups represented in the archaeological record may well have pursued a foraging strategy at least part of the time. A model such as Binford's, which contrasts two extreme subsistence and mobility/settlement strategies—foraging and collecting—is not meant to reflect the real world as much as to provide a basis for predictions. All actual human strategies should fall somewhere between these two extremes. Among groups that depend more heavily on logistically organized collecting strategies, there are definite nodes or foci in the landscape that are repetitively used for the same or different purposes. Even among near-classic foragers, such as the Bushmen described by Yellen (1976), some camps can be seen to be resettled even within the short span of ethnographic time.

Most North American prehistoric and ethnohistorically recorded hunters and gatherers could be expected to employ subsistence strategies more closely resembling the collecting portion of Binford's model and thus to exhibit a logistic mobility/settlement pattern. For example, most Shoshone groups of the Great Basin, who exploited only wild foods even ethnohistorically, occupied a number of functionally differentiated types of camps. Four major food sources were exploited: Indian ricegrass seeds, piñon nuts, jackrabbits, and antelope (Powell 1980). Winter villages served as residential bases, and foraging for seeds and rabbits took place near these camps; in addition, at least two types of special-purpose camps were occupied. Piñon camps, which were reused when the nuts were locally available, were occupied by one or more families for periods ranging from 2 weeks to several months. Antelope camps were also reused, although only about once every 12 years owing to pressure on antelope populations. When these antelope camps were in use, however, they were occupied by a large population consisting of many residential groups, and they were spatially quite extensive. The Shoshone antelope drive camp is a good example of a location being chosen not on the basis of "multivariate"

determinants but instead because of the presence of a single resource. As Thomas (1983:79) notes, at antelope camps "the short-term gain[s] of high-bulk animal procurement temporarily offset the high costs of transporting essentials such as firewood and water."

The pastoralist Navajo also exhibit differentiated use of locations within a home range centering on a permanent camp. Some of these functionally specific locations are used only once or infrequently (temporary windbreaks, tent locations), but many more are revisited regularly (e.g., stock shelters, storage features, dumps, antelope hunting corrals, sweathouses; Kelley et al. 1982). Although commonly characterized as pastoralists, the Navajo also grow crops, and they maintain agricultural fieldhouses when the distance from the permanent camp to the field is greater than ca. 3.2 km (Russell 1978). Some of these fieldhouses are occupied for the entire agricultural season, and commonly they are reoccupied from year to year.

As people become more intensively agricultural and residentially sedentary, their logistic use of nonresidential locations may actually be greater than that of hunter-gatherers. But because there is little residential relocation, these special-purpose locations are used for more or less the same set of functions, although not necessarily all at the same time. Among Pueblo agriculturalists, both living and prehistoric, special-purpose sites have often been lumped under the rubric of "fieldhouses," although they may have had many functions, including agricultural camps, lookouts, hunters' camps, and storage facilities (McAllister and Plog 1978; Moore 1978). Mesoamerican analogies suggest that small fieldhouse locations originally occupied for purposes of tending agricultural fields may grow into larger residential villages through time (Fish and Fish 1978). Ellis (1978) observes that among the New Mexico Pueblos most fieldhouses belong to single individuals and thus represent recurring occupations for only a generation. She also notes that these structures are used not only while fields are being tended but also for "vacations."

Implications of Variations in Settlement/Mobility Patterns for the Archaeological Record

Binford's model of hunter-gatherer subsistence strategies and their concomitant settlement/mobility organizations has suggested two polar extremes, that of subsistence generalists with a foraging pattern of spatial use, and that of specialist collectors whose use of space is logistically organized. Ethnographic and ethnoarchaeological documentation provides support for the conceptual validity of both of these patterns and also suggests that most groups occupy a position somewhere between these extremes. Prehistoric systems also can be expected to fall somewhere on this continuum—in other words, some aspects of their use of space will be continuous and other aspects will result in the reuse of places for the same or different functions.

What are some of the implications of these patterns for the formation of the archaeological record, particularly with respect to predictive modeling? A first

obvious implication is that within any single organizational system there should be a number of different site types with functionally and formally different contents. The determinants of the placement of these different site types vary, with some site locations (those of residential sites in logistic systems, for instance) being compromises among the locations of known resources (i.e., determined and multivariate). Some site types, for instance residential bases and locations in foraging systems, will be far less predictably located on the basis of correlations with resource locations. Still another class of sites, special-use camps within logistic systems, may be locationally quite dependent upon the occurrence of a single resource and independent of the occurrence of other resources. In order to predict the locations of special-use sites, one would need to know just what resources were being exploited at and around them. The archaeological dilemma about the function of "field-houses" illustrates that it may be very difficult to determine the specific use of places by simply inspecting those sites. Nonetheless, all of the site types that constitute a settlement/mobility system are integral participants in the overall organization of that system, and they must be understood before the locations of other components of that system can be predicted. Another implication of these patterns of space use and reuse is that a large and important portion of the archaeological record may be relatively continuous across the landscape, difficult to discover using current survey methods owing to low density of discarded materials, and very hard to talk about in terms of any equivalency between perceived clusters of materials (sites) and past behavioral episodes.

The reuse of places through time also raises questions about the practice of equating clusters of materials with sites, at least insofar as sites are automatically interpreted episodically and as having locations that are predictable on the basis of their proximity to important resources. Moreover, site size as a functionally discriminating factor may be skewed by the reuse or lack of reuse of structures or of the places where previous structures had been. In the residential camps of the northern Ute, for example, menstruating women were required to build a new menstrual hut each month; these were similar to family shelters in size and functional characteristics, having internal hearths and activity areas, and they did not occupy areas where previous menstrual huts had been built (Smith 1974). If a hypothetical northern Ute residential camp were occupied by an extended family including eight adult females, half of whom were pregnant at any given time, approximately 68 new menstrual huts would be constructed each year. If each old hut structure remained visible for 50 years, as some taphonomic studies indicate might be possible, and if the camps were continuously occupied over these years, this single Ute camp would accrue 3400 menstrual hut locations. What would normally be classed as a very large site may actually be the remains of multiple reoccupations of a single location by a relatively small group.

An exciting account by John Wesley Powell, an ethnographer who worked with the southern Numa (Ute) for two decades beginning in the 1860s, illustrates the consequences of reuse of the same general area, but not of the exact spots where structures had previously been built, at residential camps.

> It is very rare that a site for a camp is occupied a second time and though they all go again year after year to camp near the same spring or small stream they invariably seek a new site for their bivouacs each time. When they leave a camp their bivouacs are not destroyed and so on coming to a customary camping place of the Utes, it gives the appearance of having been occupied by a very large tribe, and persons are easily led to suppose that thousands have been encamped there when in fact perhaps a small tribe of a dozen families have been the only persons who have occupied the ground for many years [Fowler and Fowler 1971:52].

The nature of site patterning and the appearance and visibility of archaeological sites are seldom determined solely by the activities carried out during a single occupational episode. The archaeological record is instead created by the repetitive superposition of materials resulting from adjustments of human systems to their landscape through mobility. All components of these systems must be located, studied, and understood through the explanatory process before any can be successfully predicted.

TECHNOLOGICAL STRATEGIES, DISCARD BEHAVIOR, AND THE ARCHAEOLOGICAL RECORD

Analyzing the differences and similarities among and within collections of cultural materials that are found at places—that is, assemblage variability—is often thought of as something to be done in the future, after the cultural resource manager's work has ensured the protection of significant sites. Unfortunately, this cannot be the case in any program directed toward predicting the locations or other characteristics of sites and resources. In order to understand the workings of past systems and the mechanisms behind the spatial organization of activities, we must be able to tell the parts of systems from one another. In this section we will suggest that the component parts of human systems can be identified on the basis of the tools and other materials discarded, combined with information about the organization of technology.

Modeling Technological Organization

Ongoing cultural systems occupy a set of functionally and spatially differentiated places. If we study these places simply by grouping together sites that are similar, we cannot hope to understand the system as a coherent whole. In order to understand past systems we must find a way to group together the *different* parts of a *single* cultural system or type of adaptation. Such parts of the cultural system may occur in the form of clumped distributions of artifacts and features resulting from a single or from multiple occupations. Assemblages of artifacts resulting from different functional activities and formed at the same or different times may overlap wholly or partially in space. In other circumstances artifactual materials may be relatively sparsely and continuously scattered over large areas as a result of extensive foraging.

On the organizational level, it is clear that the archaeological record is not directly or simply equivalent to activity areas or sets. It is accretional rather than episodic, whether it is of a continuous nature or concentrated into clusters. It is necessary to sort out the overlapping, accretional sets of artifacts and features before functions and roles in the organizational system can be assigned to what we see in the archaeological record and before we can approach any sort of locational predictions. A consideration of intersite and intrasite assemblage variability is a necessary starting point.

Assemblage variability can be predicted through reference to the model of subsistence, settlement, and mobility detailed above. It should again be emphasized that models are heuristic theoretical constructs that permit us to consider the range of strategies that human groups might follow and to predict the expected results of these strategies. Models allow prediction of consequences; if these predictions are confirmed, this tends to validate the usefulness of the model. Consequences are predicted from the model through the use of middle-range theory (Figure 4.1).

Curated vs Expedient Technology

As an example of a middle-range theoretical concept with great potential for tying together the dynamic organization of past human systems and the static contemporary archaeological record, consider the distinction between curated and expedient tools (Binford 1976, 1979). Expedient tools are those that are manufactured in the immediate context of their use when the circumstances that require them arise. Examples of expedient tools are rare in today's manufactured technology, but we all use bent coathangers to open locked automobile doors or convenient sticks to chase frightening dogs. In systemic terms, the use of expedient technology would be expected to be greatest in organizational systems geared toward an encounter strategy—that is, foraging systems. In the environments that favor such a strategy, there is an equal chance of coming across a wide variety of resources; there is no need for the participants in such a system to even attempt to predict what they will find. Other things (such as material availability) being equal, it might well be most efficient for these people to manufacture tools on the spot to meet specific situations as they are encountered.

In curated technologies, on the other hand, the tools that are employed are planned to fit specific uses that have been anticipated (Binford 1976). This is an efficient strategy in environments where the occurrence of resources is predictable, and in organizational systems that focus on specialized resources. Collecting strategies featuring a logistic organization of mobility—dispatching of special task groups to procure selected resources—are most likely to exhibit curated technologies.

As in any modeling effort, of course, these two technological extremes are theoretical constructs. Actual technologies employed within a system can be expected to be a combination of the two. For instance, foraging people may produce

and use general-purpose curated tools in addition to manufacturing situational tools. It is probable that the participants in logistically organized systems will encounter unplanned situations that require the fabrication of expedient tools or the modification of tools with planned uses into tools with new uses. One characteristic of curated components of a technology is that they are often the result of *staged manufacture* employed in the face of time stress (Torrence 1983). Time stress occurs when resources are clumped or concentrated in space (which requires a focus on specific resources to consumer needs) and in time (which requires highly efficient, specialized tools). Since collecting resources in such an environment must be done in short time periods, there is plenty of time to work on tools; high energy expenditures in tool design, manufacture, and maintenance assume technological efficiency. Typically, tools are manufactured and maintained in a staged manner, with stages taking place not only at residences but also at special-purpose locations occupied on the way to and from locations of time-stressed resource procurement. Staged manufacture, resource specialization or focalization, and the use of special-purpose locations are characteristic of logistically organized groups.

Foraging groups are characterized by relatively broad-spectrum resource bases—they are generalists in that they exploit a large number of resources, at relatively low levels, within a foraging radius even over short time periods. While specialists in simple environments must obtain most of their resources during very short time periods, this is not the case for foraging generalists, who obtain food slowly and constantly. In such a generalist scenario there is neither the need nor the opportunity for staged manufacture. If technological components are curated, they are manufactured, maintained, and discarded at residential bases. Expedient portions of a foraging group's technology will be discarded continuously throughout the foraging radius.

These crosscutting but definitely not independent middle-range dimensions of variability—collecting vs foraging, resource specialization vs generalization, and tool curation vs expediency—are important in a discussion of predictive modeling in that they have different implications in terms of the location of the manufacture, maintenance, and discard of tools and hence the formation of the archaeological record. Expedient tools are manufactured where they are needed, and they are also discarded there. In this strategy, the occurrence of expedient tools is isomorphic with the activities for which they were used, and the energy put into these objects is low; they exhibit little in the way of formalization or style. Most expedient tools probably do not look much like tools at all and are therefore either exempted from analysis by many archaeologists as "undiagnostic" or included in the category of debitage.

Curated tools, on the other hand, are rarely either manufactured or discarded in the context of their immediate use. Tools intended for use during the mobile activities of special task groups are most likely to be manufactured at residential basecamps (Binford 1980) for anticipated uses away from those camps. Curated tools, designed to be used for some time, will be more durable than those made expediently for immediate discard, although this may not be morphologically

obvious. Characteristically, however, curated tools are of a compound or complex nature (Allchin 1966; Oswalt 1973), having hafted components or multiple parts. These characteristics help to ensure that a curated tool will not be "used up" at the locus of its use but rather will be brought back to the residential base for rejuvenation or other maintenance. Under such a curated technology, both manufacturing debris and broken portions of curated tools will be found at the residential site and not at the places where the tools were used. The only curated tools (as opposed to "site furniture," such as metates) that should be found at the place where they were used are those that were lost and not recovered, for instance, unrecovered projectile points.

Expectations about the presence of discarded tools and debris associated with tool manufacture in the archaeological record can be generated from the above assumptions. Under a foraging strategy, there are two situations in which discard should take place: at the residential basecamp and at the location. Manufacture and discard of expedient tools would be expected to take place at both of these loci, with the implements being discarded where they were used. Groups using a foraging strategy should exhibit major variations in mobility and group size and composition during the year or from year to year in response to short-term variations in the environment (Binford 1980). This leads to the expectation that the activities performed at foraging sites of either type could be quite diverse and could vary with time. Since over the long term, at least, campsites would not be chosen with regard to the placement of previous camps or locations, this diverse archaeological record, particularly those assemblages derived from locations, would tend to be relatively continuous over the landscape, given long-term use. Under a foraging strategy, variability in residential site assemblages is the result of differences in seasonal scheduling of activities and in duration of occupation. In such systems there is a pattern of increasing assemblage diversity with increasing site size, as noted by Yellen (1976). Among groups practicing a foraging strategy, therefore, the nonresidential use of the foraging radius leaves *nonsite* archaeological remains that are just as important for archaeologists attempting to predict the operation of these past systems as are the more clustered and visible materials that are usually called sites. This problem of continuous distributions will be discussed at greater length later in this chapter. It is quite likely that some components of *all* human systems leave dispersed archaeological remains with low visibility, and these remains must be studied and understood before the mechanisms behind the placement of activities in systems can be explained and used to predict the locations of those activities accurately. The record left by expedient activities may be far more easily understood than that of the more logistically organized portions of past systems.

Under a logistically organized system the nature of the intra- and interassemblage variability can be expected to be very different from that predicted for foraging systems. Collectors use specially organized, highly mobile task groups to accommodate situations in which consumers are near one or more critical resources but distant from others. In addition to residential basecamps, these groups also utilize field camps, stations, caches, and other places for specific functions. Field

camps under such systems probably outnumber residential camps by as much as 4:1 (Judge 1973). Since these camps can be occupied for long periods and/or be the sites of intensive processing, they may become as large and visible, archaeologically, as residential bases (Binford 1980). As noted above, groups organized under a collecting strategy will be likely to employ a curated technology to some extent, given their high levels of mobility and activity planning. Discard of those curated tools that are employed primarily away from the residence rarely takes place at the locus of their use. Collecting strategies are based upon prediction or planning and should be expected to occur for the most part in the most predictable environments. This means that places where archaeologists today should best be able to predict the locations of sites on the basis of resource distributions will harbor assemblages that are unlikely to reflect the activities that took place there, since they will have less functional correspondence with the "resources" that are used as independent variables in predicting them.

The argument might be made that it is not necessary to know the functions of sites to be able to predict their occurrence—that using proxy indicators that can be measured in the environmental today and that "predict" the occurrence of sites empirically works just as well. This may be true in certain situations, but proxy indicators should not be expected to occur isomorphically with the reasons that activities took place at certain locations in the past in all cases. It is the *mechanisms* behind the placement of activities in space and their resulting archaeological record that must be understood in order to successfully predict the occurrence of activity loci.

The Reuse of Places and Intra-Assemblage Variability

Attempts to predict the occurrence of sites that result from the operation of logistically organized systems are further complicated because places are reused for different purposes, so that many different combinations of activities may take place at a single site. For instance, a place might be used as a residential base for several months and thus contain tool manufacturing and maintenance-related debris. If the site were subsequently used as a field camp, the discarded materials from this second use may not faithfully represent activities that actually occurred there. A wide range of technological variability of specific and easily differentiated types can be expected in the archaeological record produced by a collecting-based systemic organization. Investment in such facilities as structures for shelter or storage, caching of items to be used later at the site, and other cultural "improvements" of a place would also be expected at reused places under such a system. This means that differential site function in a logistically organized, collecting system *might not be obvious* on the basis of either site size or site contents. Indeed, as Thomas (1983:80) points out,

> it is extremely difficult to distinguish field camps from base camps in the archaeological record. There are behavioral differences to be sure, but these differences are commonly subtle and off-the-cuff field designations should always be mistrusted.

Interassemblage Variability and Mobility

The variability among assemblages at different sites that result from the operation of a single system—that is, *interassemblage* variability—is the result of the overlay of an organized series of events. The nature of assemblages that result when cultural events interact differentially with natural events has been discussed in terms of "grain size" by Binford (1980:17). Coarse-grained assemblages are the cumulative product of events spanning relatively large time periods, for instance several months or a year. Fine-grained assemblages accumulate over a short period of time. The finer the assemblage grain, the greater the probable content variability *among* assemblages, because there is less chance that the total range of activities that occur under that system will be found there. The main factor responsible for grain size is mobility, but this relationship is far from simple or linear. In a foraging group, residential mobility would be expected to be highest in the least diverse, least seasonal, and least predictable environments, resulting in an increase in interassemblage variability. Under logistic strategies, residential mobility goes down, so coarser-grained assemblages would be expected in residential sites; the more mobile logistic components would, however, be finer grained than the residential sites and would thus, as a class, exhibit more interassemblage variability.

The Explanation of Intra- and Interassemblage Variability

Two major expectations concerning the relationship between assemblage variability and differing degrees of residential vs logistic mobility have been discussed above. One expectation is that under increasing logistic mobility the effects of curation and the reuse of places will make it increasingly difficult to postulate the functions of sites or to predict their occurrence in terms of association with particular resources. The other expectation is that under increasing logistic mobility there will be increased interassemblage variability, both between residential basecamps and special-task locations and among different special-task locations as well. The archaeological record in this latter case may appear as a series of sites that are relatively uniform in size, visibility, and contents in terms of structures or facilities but contain assemblages that are strikingly different in terms of the formal attributes of their constituents or at least some of their constituents.

One of the ways of explaining an archaeological record like that described above is in terms of separate technical or cultural traditions, an approach that has been dominant in American archaeology since the science's beginnings (Willey and Sabloff 1974). This approach, which has been referred to as the Kriegerian method (Binford and Sabloff 1982:143), defines culture types as collections of formally similar properties or attributes of cultural materials that are spatially coherent. Data collected and interpreted using this approach pose serious problems for archaeologists and cultural resource managers who wish to understand the operation of past human systems and the mechanisms behind the archaeological record, yet such an understanding is critical to successful prediction of the locations of archaeological

materials. A Kriegerian, culture-type approach to assemblage variability virtually assures that the differentiated components of systems will be treated as separate cultures or traditions, making it impossible to consider them as parts of an integrated whole. And since the components of a system do operate in an integrated manner, their locations are just as dependent upon the nature and locations of the *other* components as upon environmental or other factors. Successful prediction is necessarily based upon the recognition and sorting out of the complementary components of systems. Unfortunately, this is something that archaeologists cannot do at present, although attempts toward this goal will be discussed later in this chapter.

Technological vs Ecosystems Organization

The practice of grouping assemblages on the basis of formal similarity encourages an emphasis on empirical correlations between assemblages (site types or culture types) and environmental variables, a practice that is the hallmark of present-day prediction attempts. Mazel and Parkington (1981) suggest that a more productive approach might consist of regional studies of the interrelationships among tools, sets of tools, and resources. These interrelationships are controlled, they feel, primarily by the spatial patterning of resources (rather than simply by their location) and by the ways in which resource patterning compares with the spatial patterning of human mobility within a system. In other words, prediction might be based not only on an understanding of human systems but on knowledge of ecosystems as well. Ecosystem variables include the *patterning* of resources in time and space and such qualities as environmental diversity and equability. The effects of ecosystemic spatial and temporal structures on the predictive effort will be discussed later in this chapter.

Selection of the cultural variables against which to compare ecosystem variables may be one of the most difficult tasks presently before the archaeologist. It will require very different approaches to sampling, survey, and data collection, recording, and analysis than are used in cultural resources management today. The assemblages that constitute sites must be understood in their entirety—undiagnostic artifacts as well as diagnostic ones. One new approach, a nonsite or distributional archaeological survey method, was recently tested by the Bureau of Land Management in New Mexico. This project will be discussed in a later section of this chapter.

From a systems perspective it is clear that, at least under certain types of mobility and technological organization, the contemporaneous technological "traditions" often identified in the archaeological record are actually functionally different parts of the same system. Most of the archaeological record in any one place may consist of the remains of different portions of an essentially similar system—remains that have been deposited over very long periods of time. The archaeological record is not directly explainable in terms of episodic behavior; rather,

> a detailed consideration of the factors that differentially condition *long-term* range occupancy or positioning in macro-geographical terms is needed before we can realistically begin to develop a comprehension of . . . subsistence-settlement behavior. The latter is of course necessary to an understanding of archaeological site patterning [Binford 1980:19; emphasis original].

NATURAL FORMATION PROCESSES AND THE ARCHAEOLOGICAL RECORD

The complex patterning of cultural materials across space is a result of human mobility, the spatial patterning of different economic activities, the redundancy in economic activities across the landscape, and differences in the locus of artifact discard vs that of artifact use. In most cases, this patterning of discarded material undergoes additional changes before it is discovered and interpreted by the archaeologist (Schiffer 1972, 1983). Processes affecting the deposition, accumulation, preservation, disturbance, and exposure of the materials that make up the archaeological record have been much investigated in recent years, largely due to such interdisciplinary influences as the study of taphonomy of culturally utilized or modified organic and inorganic materials (Behrensmeyer and Hill 1980; Brain 1967a, 1967b, 1969, 1981; Gifford 1977a, 1977b, 1980, 1981; Gifford and Behrensmeyer 1977) and geoarchaeology (Butzer 1977, 1982; Gladfelter 1977).

Deposition: The Coincidence of Natural and Cultural Events

Cultural materials enter the archaeological record through deposition, during which process they are buried or otherwise preserved. Although depositional processes may be cultural, in most cases they are natural, consisting of aeolian, fluvial, lacustrine, or residual aggradation. These natural processes of deposition may or may not coincide with episodes of cultural discard. Materials discarded as the result of an occupation or activity might lie on the surface for long periods (in fact, "forever") without being buried, or they may be quickly buried even as they are discarded. Materials buried in layers or "levels" are thus not necessarily or even not often expected to be the result of single occupational episodes. The nature of the deposited archaeological record is controlled by the periodicity or "tempo" (Binford 1982:16) of occupation or use of a place and by the relationship between this occupational periodicity and the periodicity of depositional processes. If the periodicity of discard is the same as the periodicity of natural occurrences—for instance, floods—that incorporate these the artifacts into sediments, then a regularly stratified archaeological record will result. If discard occurs more often than the natural encapsulating events, however, cultural materials resulting from multiple behavioral episodes—multiple activity sets, in Carr's (1984:113) terms—will be incorporated into the same geomorphic stratum.

In situations such as the complete radius leapfrog pattern of residential mobility, for instance, in which certain logistic sites may be reoccupied or reused for

different activities within a short period of time, one might expect that episodes of discard would occur more frequently than episodes of deposition. This would result in single-layer assemblages, or what Carr (1984:114) calls depositional sets, composed of materials from more than one occupation or function. The nature of the deposited archaeological record is determined not only by the organization of the cultural system but by interactions between the organizational system and depositional processes. This poses another set of problems for the archaeologist, since "demonstrably associated things may never have occurred together as an organized body of material during any given occupation" (Binford 1982:17–18).

Postdepositional Processes

Another set of processes affecting the ultimate nature of the archaeological record can be thought of as postdepositional, occurring after the discard of cultural materials. Generally, almost any process that disturbs or acts upon the surface of the earth and subsurface deposits also acts upon archaeological materials. Such biological processes as faunalturbation and floralturbation (Wood and Johnson 1978), caused by burrowing, trampling, and root-heave, can modify the original distribution of cultural materials. Chemical and physical processes that affect the archaeological record include freezing and thawing cycles; mass wasting (gravitational forces); the growth and wasting of salt crystalline structures; the swelling and shrinking of clays; volcanism and tectonism; disturbances caused by the action of gas, air, wind, and water; and pedogenesis.

A somewhat different taxonomy of the postdepositional processes acting on the archaeological record is advanced by Foley, who presents five sets of processes responsible for burial, movement, destruction, exposure, and "small-scale oscillation" (1981a:167) of archaeological materials. Discarded artifacts enter the archaeological record through burial by cultural or natural agencies; once assemblages are buried they may remain in place or they may be moved through stream action, sediment movement, faulting, or mass wasting. At the same time, certain materials may or may not be destroyed by physical and chemical agencies while in or on the ground. Small-scale oscillation processes include animal burrowing, human disturbances, root action, and water or wind action; these forces may alter the position of components of the archaeological record slightly but presumably do not totally disarrange it. Exposure of the archaeological record to water or wind erosion, tectonic activity, or human disturbance may alter the distribution of the archaeological materials as well as make them visible.

Just as variations in the coincidence of episodes of discard vs episodes of deposition or burial can create either well-segregated assemblages or palimpsests (that is, artifact distributions resulting from the overlay of many separate behavioral episodes and the action of postdepositional processes), exposure and reburial can also introduce complexities in archaeological patterning. These processes are rarely simply gravitational; they usually include some lateral component and therefore are

influenced by small variations in topography. Exposure and redeposition are often highly localized; deposits from separate occupations may be mixed in one area while a few meters away they will be separately stratified. Controlling for the complexities caused by differential deposition, exposure, and reburial of artifacts may be one of the most difficult and yet necessary tasks facing the archaeologist. Whatever the scale at which patterning in the archaeological record is being analyzed, the microtopography and geomorphological activity of surfaces must be examined in more detail than that afforded by most generally available topographic or surface unit maps.

The Scale of Depositional and Postdepositional Processes

Natural depositional and postdepositional processes are not necessarily or even often controlled by the factors that caused prehistoric people to visit and use an area. Depositional and postdepositional processes are localized and patterned on a small scale. Rarely, then, will the actions and results of these natural processes be spatially congruent with activity areas or assumed sites. Instead, their effects serve to remove the archaeological record yet further from past behavior and the organization of human systems.

This is not to say that natural processes necessarily render the archaeological record useless or uninterpretable. It is common in contemporary archaeology to view postdepositional processes as "bad," as making the archaeological record unusable or of diminished research potential. This probably arises from the seemingly popular belief that postdepositional processes are random in their operation (Bowers et al. 1983; Kirkby and Kirkby 1976). Almost all modern survey forms have a space for an assessment of a site's integrity; if the site is disturbed, it is too often classed as being of limited utility to science and therefore of diminished significance. Such an assessment ignores the fact that all archaeological materials, whether from "sealed" sites or lying on the surface, have been affected by natural processes. Depositional and postdepositional processes are not random in nature; in order to assess their effect on our data, however, we must study and understand these processes so that we can predict their distribution and impact. Any prediction of the occurrence of archaeological materials *must* incorporate a full consideration of the effects of depositional and postdepositional processes as intervening factors between the operation of past human systems and the archaeological record.

This is necessary because the effects of postdepositional processes on what we see as the archaeological record may be far greater than we intuitively recognize. They not only disarrange flakes and tools but in fact are almost totally responsible for most of what archaeologists actually *see* during surface survey. If the physical extent of behavioral events that result in discarded materials are of the same general range of spatial scales as the depositional and postdepositional processes, then there is some chance that entire sites will be exposed to the archaeologist's view. Unfortunately, it is almost inconceivable that this will be the case. The material record will almost certainly be acted upon by a series of partially overlapping

depositional and postdepositional processes of widely varying scales. These processes will combine the products of behavioral episodes; blur or sharpen (and in fact probably often *create*) their apparent boundaries; and differentially affect the placement of artifacts, depending on their sizes and shapes. These effects are all-important, for they determine where we see sites and what these sites look like. They also may be responsible for the fact that we think we see "sites" at all in many places. These processes must surely be determinable and predictable. The natural processes that intervene between the archaeological record and our knowledge of the past must be understood before predictive modeling can become an operational tool for cultural resource management. This task is discussed and illustrated at length in Chapter 9 of this volume, which deals with remote sensing and predictive modeling.

The Usefulness and Integrity of Surface Remains

Recently Lewarch and O'Brien (1981) argued that surface assemblages can be used to answer archaeological questions because comparable processes affect the patterning of artifacts in both surface and subsurface archaeological assemblages. A more realistic way to phrase this might be that archaeologists should be aware that natural and cultural processes can affect subsurface or "sealed" archaeological patterning just as strongly as they affect surface materials, so no automatic assumptions of total contextual integrity should be made for any observed archaeological patterning.

There are important pragmatic reasons for developing methods to measure the patterning and content of archaeological surface remains and for using such data to answer archaeological questions. Of these, the most relevant to the present volume is that the depositional processes that seal and protect cultural materials after their discard usually render these materials invisible to the archaeologist, even when such sophisticated and often expensive techniques as underground radar, proton magnetometry, resistivity measurement, and the like are used to search for them. For practical purposes, most buried archaeological materials are unknown and of no value to the archaeologist until they are exposed. Another reason for paying attention to surface assemblages is that the contexts in which stratified deposition and burial are most dependable and regular, and in which archaeologists most often look for and find buried materials, may be the result of only very limited or specialized portions of the cultural systems. For instance, while cave sites contain well-segregated and well-preserved cultural strata, such sites might have been occupied only when the shelter they afforded was necessary, or they may have been used only for a specific set of purposes. Most of the components of the cultural system may have involved the use of open situations that would be more likely to be buried and reexposed, or perhaps not buried at all. Thus, in the archaeological record these components would be represented only by surface assemblages.

Possibly the best reason for using surface archaeological assemblages, however, is that such data can be collected quickly, accurately, and cost-effectively, and they

yield a high return in the form of information that can be used to test models of human systems organization. In order for us to use this information, however, it is imperative that surface archaeological data be discovered, measured, and analyzed in ways that are consistent with their nature and with the nature of the organizational processes that we wish to explain, as documented in the final section of this chapter.

Natural Processes and "Independent Environmental Variables"

The importance to predictive modeling of an understanding of postdepositional processes becomes clear if we consider the "independent variables" frequently discussed by archaeologists involved in locational predictive modeling. These independent variables are the noncultural aspects of the total environment that correlate with site locations. Under an empirical framework these variables are used to "predict" (project) site locations. Commonly used independent variables include soil association, slope, elevation and/or variation in elevation, topographic aspect, vegetation, distance to water sources and their nature, and various specific landform associations (Chapter 9). It is almost always explicitly acknowledged that these independent variables themselves may have no causal relationship with the placement of sites; they are simply considered to be indicators. In many instances, variables may be chosen primarily because they can be taken conveniently and quickly from topographic maps so that fieldwork is not required; some of the pitfalls of this approach will be discussed in Chapter 9.

In addition, trying to generalize about where prehistoric people lived on the basis of where we find their discarded materials circumvents the explanatory framework outlined above by equating the archaeological record with past behavior without taking intervening processes into account. Correlating environmental characteristics with the archaeological record must begin with a consideration of the natural processes that determine how we see the archaeological record. Every one of the independent variables used in empirical, correlative projections could be a successful predictor because it has relevance to *natural* depositional and postdepositional processes (and thus to the visibility of archaeological materials) rather than for any cultural reasons.

For example, archaeological materials might be found on ridge tops, in sand dunes, or near water sources because that is where they are exposed and visible today. Soil associations are taxa of different types of soils, and these differences are based largely upon varying parent materials and the time that the soil has had to develop, both of which may affect the geomorphic processes that cover or uncover artifacts. Vegetation is an obvious factor in reducing or enhancing archaeological visibility. Erosion takes place at accelerated rates on steep slopes. And any archaeologist who has tried to survey the north side of a hill in the early morning or late afternoon knows that the light there is poor; things can simply be *seen* better on south slopes. There is not a single independent variable used in current predictive

modeling attempts that might not have more to do with depositional and postdepositional processes than with anything that prehistoric people thought or did.

Predictive modeling based on correlations with these variables may actually be predicting where we see sites and may have very little to do with how people behaved or how their systems worked. This is not to say that the archaeological record has no systemic, behavioral determinants; previous sections of this chapter have emphasized that it does. The point is that natural processes are very important in determining many aspects of the nature of the archaeological record and how we, as archaeologists, can deal with it. They must be thoroughly understood before predictive modeling can become either a management or a research tool. Ways of measuring, understanding, and even predicting the effects and distribution of natural depositional and postdepositional processes will be discussed more exhaustively in Chapter 9.

ECOSYSTEMS VARIABLES AND ARCHAEOLOGICAL EXPLANATION AND MODELING

As defined in the first section of this chapter, archaeological explanation is the process of combining middle-and upper-range archaeological and anthropological theory with ecosystems theory to form models from which predictions are drawn. This process begins at the systems level, and archaeological models connect systemic human organization with predictions about the archaeological record.

Human systems obviously exist within ecosystems—they are subsets or components of ecosystems. Ultimately, the nature and predictability of human systems and their products will be related at least in part to the natural ecosystem. This is an explicit assumption in all predictive modeling or projective attempts known to the authors of this chapter. In fact, the almost universal approach for such attempts is to compare the distribution of archaeological materials with "environmental variables" that are suspected of having been important to past people: the availability or lack of water, shelter, firewood, food species, lookouts, south-facing slopes, etc.

This section will discuss the use of eco*system* variables rather than particularistic environmental resources in the process of archaeological explanation. Ecosystem variables have considerable explanatory power when incorporated into models of change in human systems in response to ecosystem properties; they also have implications for the ultimate "predictability" of locations of cultural resources in different ecosystemic settings. In keeping with the principle of congruence in levels of systems being compared, it is important to examine the global characteristics of the structure of the ecosystem in order to predict something about the structure of the human organizational system inhabiting it (Figure 4.4). On a lower level, the spatial and temporal distribution of that environmental structure is important for predicting the spatial and temporal distribution of the human system exploiting it. At a still lower-order level in both systems, it is important to be able to characterize

	ECOSYSTEMS		SETTLEMENT SYSTEMS		ARCHAEOLOGICAL MANIFESTATIONS
Global characteristics	Net primary productivity Potential evapotranspiration Degree of seasonality Accessibility of resources	affects →	Population size and density Resource mix Degree of mobility	affects →	Abundance of archaeological materials or sites Assemblage grain Number of site types
	Components		**Components**		
Characteristics of spatialized system components	Spatial patchiness Patch diversity and overlap Temporal predictability of patch Ruggedness of terrain	affects →	Concentration of activities within the landscape Diversity of activities at loci Degree of site reuse	affects →	Concentration of materials in space Diversity of depositional sets (assemblages)
	Components		**Components**		
Elements of these spatialized system components	Spatial and temporal distribution of biotic species Abiotic components of the environment	affects →	Spatial and temporal distribution of activities Episodic behavior	affects →	Artifacts and activity sets

Figure 4.4. Parallel levels of ecosystems and settlement systems and their archaeological manifestations. Both ecosystems and settlement systems are components of a single regional system at a higher level. The causal arrows shown are not meant to be exhaustive.

the distribution of particular resources in order to predict the location of specific prehistoric activities and their archaeological manifestations. The extent to which this is possible, however, will depend on higher-order characteristics of both systems.

Human Systems Within Ecosystems

Ecosystems are composed of individuals enmeshed in populations, interacting with other populations in communities. Ellen (1982:74) has provided us with a useful modern definition of the ecosystem as

> a relatively stable set of organic relationships in which energy, material, and information are in continuous circulation, and in which all processes are seen in terms of their system-wide repercussions. *Specific* changes, which may theoretically begin anywhere in the system, trigger adjustment and re-adaptation among the other elements. . . . *Systemic* changes take place slowly through conjoint evolution that is biological, chemical, and physical.

The ecosystem composed of these interacting communities is another example of a general living system and likewise exhibits a mixture of predetermined behavior and free systems dynamics, as discussed earlier in this chapter and in Buechner (1971:45). The species composition of particular locations in a forest, for example, is always changing in response to fire or other perturbations, although species composition and dominance in the larger forest may remain relatively stable. Species composition in seral (i.e., successional) communities varies according to both random and predetermined processes (Buechner 1971:52–53).

On an abstract systems level, a number of relationships between ecosystemic characteristics and aspects of settlement systems have been demonstrated or suggested. Binford (1980) remarked upon the increasing importance of both logistic mobility (collecting) and storage among hunter-gatherers in environments with increasing seasonality. He notes that foragers, who practice little storage or logistic collecting, tend to move from the center of one resource area to the center of the next. Kelly (1983) has argued that the resource "accessibility" (the amount of time and effort required to extract resources from an environment) of plants can roughly be estimated by dividing the net above-ground primary productivity of an environment by its primary (plant) biomass; animal accessibility is roughly measured by dividing secondary biomass by primary biomass. (Net primary productivity is the rate of increase over some unit of time in biomass, usually measured in calories.) Kelly finds that as resource accessibility measured in this way decreases, residential mobility increases. Low resource accessibility and high residential mobility are, in turn, correlated with short distances between sequential residential bases, as is typical for foragers in the tropical rainforest.

Some Factors Affecting the Predictability and Location of Human Use of Space

An appreciation of mobility is vital to our understanding of how the archaeological record is formed. The causes and consequences of mobility are only part of what we need to know, however, in order to predict the locations and characteristics of past human behavior. In particular, we need to consider the middle and lowest levels in the systems of hierarchy shown in Figure 4.4, both for ecosystems and cultural systems. The middle level for ecosystems consists of information about the spatial and temporal structure of the ecosystem in some region of interest. The lowest level consists of information about how the distributions of specific resources make up the patches and about specific environmental features (soils, landforms, etc.) of the landscape.

Of the many kinds of knowledge that might improve our ability to understand settlement systems and to estimate *how well* site locations may be predicted, three dimensions of variability are most important: the *temporal* and *spatial variability* in resource availability and the degree of economic *intensification* of the people exploiting those resources. We will first define these three dimensions of variability and then explore the effects of each variable on settlement systems; each variable will first be discussed as if it were possible to hold the other two constant. Finally, we will give some concrete examples of how these three independent dimensions of variability can be used to characterize various settlement systems and environments in terms of the likely success of the prediction of settlement locations.

Spatial heterogeneity in the landscape is called *patchiness*, a term that is not readily quantifiable but refers to significant spatial discontinuities in the distribution of populations or communities. Intuitively, it is the opposite of homogeneity; although all ecosystems are patchy at some scale, the relative homogeneity of the tropical rainforest, for example, distinguishes it from the relative patchiness of a semiarid landscape. Patchiness encompasses aspects of environmental variability that are measurable, including the size and size distribution of patch types, the relative differences between patches and their surroundings, and so forth (Winterhalder 1980:153).

Three terms are especially useful for describing the temporal distribution of resources (Colwell 1974; Winterhalder 1980:162–163). *Constancy* is a measure of the degree to which a resource is continually available. Rainfall has a high constancy in tropical rainforests but a low constancy in most areas of the North American Southwest. *Contingency* is a measure of the degree to which the availability of a particular resource can be accurately predicted based on the season, without the need for monitoring that resource. In many areas of the Pacific Northwest, anadromous fish runs have high-contingency predictability even though they are not constant. Perfect temporal *predictability* for a resource can be due to perfect constancy, perfect contingency, or a combination of the two. For example, Bella Coola, British Columbia, has moderately predictable rainfall patterns owing to relatively high constancy coupled with relatively low contingency. Acapulco, Mexico, has equally predictable rainfall as a result of low constancy coupled with high contingency (Colwell 1974:1151).

Intensification has at least two manifestations. It may refer to the process of expending increasing amounts of time or energy to realize the same level of returns, or it may describe the process through which the same amount of output is obtained from less and less land—either through increased time or labor inputs or through more efficient technology. Intensification appears to be closely related to a number of factors: increasing involvement of groups beyond the family in the regulation of production (Sahlins 1972:101–140), increasing population, increasing population density, approach to a current carrying capacity, and increasingly complex sociopolitical organization (Harris 1977), to name a few. Harris's position (1977:70) that increasing population and increasing population pressure on resources results in intensification of land and labor, which in turn causes increasing sociopolitical complexity, may be too unilinear, but the general correlation of this system of variables is clear.

Boserup (1965), Binford (1983:195–232), and many others have discussed factors that may be seen either as the causes of intensification or as its symptoms: increased population size and packing, decreased mobility, the beginnings of serious agriculture, increased sociopolitical complexity, increased importance of exchange, the rise of urbanism, and so forth. Intensification is used here simply as the name for this large system of covarying variables, organized along the lines proposed in Table 4.1. Under certain circumstances intensification may involve the adoption of agriculture (Binford 1983:205) or the development of industrialism (Wilkinson 1973).

Some of the following discussion of the effects of spatial and temporal distribution of resources and degree of intensification on human settlement systems is exploratory, and we know of little empirical proof for some of the relationships suggested. This is a starting point for further work in this direction and serves as a qualification to simple empirical correlations of the locations of sites with environmental variables.

TABLE 4.1.

Selected correlates of intensification

	Degree of Intensification			
	Low	→	→	*High*
Correlates	*Foraging**	*Collecting**	*Casual or Extensive Domestication*	*Intensive Domestication*
Modal group size	small (18-120)	moderate	large	large
Generic site types	1-2	5 or more	many	many
Residential mobility	high (15-50 moves per year)	moderate to low	low	low
Investment in facilities	low	moderate	high	very high
Storage	very little; food gathered daily	seasonal	seasonal	long-term

* All information on foraging groups and generic site-type information on collecting groups from Binford (1980)

Intensification. What is the importance of intensification for our ability to predict where sites might be located in space? What implications might it have for the value of the concept of the "site"? The least intensified hunter-gatherer economies, people practicing a foraging way of life, should exhibit high residential mobility, practice little storage, gather or hunt food almost daily, and conduct much of their hunting on an encounter basis. The Dobe !Kung and the Central Kalahari San (Tanaka 1976) provide good examples of foragers.

Although all human systems may well exhibit some foraging subsistence, and thus mobility, behavior, the purest examples of foraging should be found in tropical areas where there is relatively little seasonal pulse in the availability of resources. Ignoring for a moment the effects of such ecosystemic factors, the following observations about foraging systems in general can be made:

1. In comparison with logistically organized hunters and gatherers (collectors), foragers should exhibit low population densities and expend relatively little energy in food transport and processing for storage.

2. The tendency for foragers to move themselves to food and water, rather than vice versa, suggests that distributions of such resources may in general be good predictors of *residential* bases (if these can be distinguished in the archaeological record). As a cautionary note, however, see comments by Foley cited earlier in this chapter. Yellen (1976:52) also observes that the !Kung San in the northern Kalahari—whose site locations are heavily constrained by the availability of water—generally locate their residential bases at least one-half kilometer, and often much farther, from a water source so as not to disturb the animals that also make use of the water.

3. Unless the environment is very homogeneous, or unless there is a single resource that is overridingly critical (such as water), however, the residential bases of sequential foragers may be located with respect to *different* suites of resources, since residential bases are used for a short time.

4. The low population density of foragers suggests that there may be a low tendency toward reuse of residential bases (what Binford [1980:7] calls redundancy in the occupation of particular places) except where there are significant topographic or other constraints in the physical environment.

5. Given long-term use of an area these last two observations may mean that all favorable resource locations will be occupied. But the small group sizes, short duration of occupation, and low rates of residential reoccupation will lead to low archaeological visibility, low artifact density, and little boundedness in space, making application of the "site" concept relatively difficult and arbitrary. Groups practicing foraging also conduct activities away from their residential bases, and activities at these "locations" (Binford 1980:9) can be expected to leave only very low densities of archaeological materials that do not correspond to established notions of sites.

The logistically organized subsistence-settlement system of collectors represents an intensification compared to foraging. A landscape in which foraging is

possible should be able to support more collectors than foragers owing to the collectors' increased efficiency in exploiting spatially disjunct but temporally concurrent resources and in overriding temporal disjunctions of resources with storage. The implications of this particular intensification for hunter-gatherers have been explored earlier (see Table 4.1) and include greater reuse of some places in the landscape, but not necessarily for the same purposes; greater degree of disjunction between those places and any single "critical" resource; and a wider variety of site types, which can be expected to differ dramatically in their locational determinants.

Active management of plant and animal resources—including domestication—entails an additional intensification of the hunter-gatherer way of life. In environments where both collecting and agriculture are feasible, a particular landscape should be able to support more agriculturalists than collectors, since the former more effectively exploit the potential net primary productivity and overcome temporal discontinuities in resource availability. Although there are some climates in which storage is difficult, most domesticators of plants and animals practice more storage than hunter-gatherer groups. Increased storage may lead to increased investment in facilities and increased residential sedentism (Hitchcock and Ebert 1984).

Although agriculturalists decrease their residential mobility in comparison with most hunter-gatherer groups, their logistic mobility is not necessarily decreased; in fact, owing to the heavily altered nature of the foraging radius surrounding agriculturalist settlements (Kohler and Matthews 1988), logistic mobility may be more frequent, and encompass a wider radius, than among groups with a more mobile residential base. Among these groups, however, logistic procurement as a means for coping with resource shortages is increasingly supplemented by exchange networks involving subsistence and/or sumptuary items. (This is not to imply that such networks cannot be important to nonagriculturalists in certain circumstances, as is amply demonstrated by some Archaic period groups in eastern North America or by the trade network in the Pacific Northwest centered on The Dalles.) With increasing sedentism, trips away from the residential base are increasingly likely to emphasize interaction with other groups, rather than direct resource collection from the natural environment, as their primary goal.

In general, the effects of increasing intensification in the absence of changing ecosystems variables can be summarized as follows:

1. residential mobility tends to decrease;

2. environmental perturbation in the vicinity of residential sites tends to increase; the original environmental communities are replaced by communities at a less mature stage, with higher net primary productivity;

3. logistic mobility and its supplement or surrogate—exchange—tend to increase.

Given the increasing importance of exchange relationships as a supplement to logistic mobility for providing access to resources outside the foraging radius, the location of other groups—and other components of the settlement systems of a

single group—becomes an increasingly important consideration in the location of residential sites. The significant disruption in the foraging radius surrounding the residential sites of agriculturalists and the possible investment in facilities within this radius (irrigation systems, for example) result in considerable pressure to keep residential sites outside the foraging radii of other residential sites. On the other hand, special economic, social, or political ties with other groups may dictate that inter-residential distances not be too great.

One implication of these changes for the *visibility* of the archaeological record is that intensification should lead to increasing visibility for residential bases because of decreased seasonality of occupation, increased longevity of occupation, increased investment in storage and dwelling facilities, and increased alteration of the natural environment.

Implications of intensification for the visibility of site types other than residences are more complicated. For locations within the foraging and field radius of the residential base that have relatively stable resources, such as arable soils, location reuse may be routine, eventually resulting in high site visibility. Within their foraging or field radius, agriculturalists or intensified hunter-gatherers invest more in facilities and revisit these facilities more frequently than do groups that regularly move their residential bases long distances; this may help to explain the relatively high visibility of "fieldhouse" sites in the American Southwest. Locations where some nonrenewable or slowly renewable resource such as wood is exploited, however, may be used in a way that is not substantially different from or more visible than the way that foragers use locations away from their residential bases.

To summarize the effects of intensification for *where* sites will be located, residential sites should increasingly represent a compromise location (Figure 4.5). Either they should be located not too far from any of the resources that will be needed regularly during the increasingly long period that such sites are occupied, or they should be located near some important subset of these resources and count on kinship ties, trade, or usufruct privileges to obtain the remainder. These predictions refer to individual residential sites, since the *total set* of forager residential bases on a given landscape may be responding to as many different environmental factors as the total set of collector or agriculturalist residential bases. Within the economically acceptable zone of possible residential base locations, considerations of comfort are not insignificant for a site that may be occupied for several years, and the locations of the residential bases of other groups become an increasingly important consideration as well.

The definition of what is a suitable zone for residential sites—both economically and from the perspective of comfort—may become broader under intensification. The increasingly complex technology that accompanies intensification permits intensive use of areas that are unsuitable for occupation by people with a simpler technology. The development of irrigation, for example, makes agriculture possible in places where it could not be practiced without irrigation. Variables determining residential base location cannot be assumed to be identical for groups at different levels of intensification.

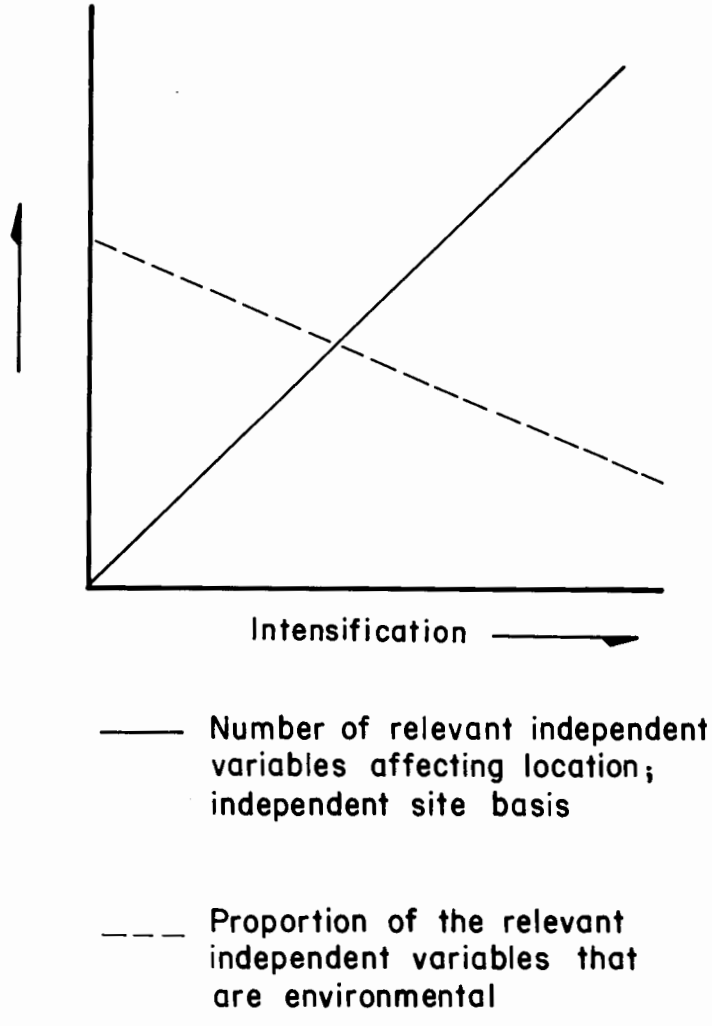


Figure 4.5. Suggested effects of increasing intensification on the location of residential sites.

Finally, site types other than residences may be located for very specific, single-resource considerations (for instance, clay or chert quarries), or they may represent compromises among several variables that are weighted rather differently than they are for residences, as is probably the case with fieldhouses.

Next, let us summarize the effects of intensification on the *predictability* of site location (that is, how strong the association between selected environmental variables and archaeological materials should be). The increased population packing under which intensification is expected to take place may mean that a smaller number of the places in the landscape that fulfill the requirements for use or settlement will remain unused; in a fully packed landscape, all suitable locations

may be used. This should make prediction easier in the limited sense that it should lower the frequency of wrong predictions about where sites are (Figure 4.6). Perhaps more important, residential bases for collectors or agriculturalists should all have similar environmental determinants within a particular settlement system, whereas forager residential bases within a single settlement system may have quite different determinants. The prediction that a single set of environmental determinants will apply to all residential bases for agriculturalists within a single settlement system is weakened, however, by the tendency for exchange to allow communities to occupy locations with access to complementary rather than redundant resources.

The implications tend to complicate inferential locational modeling. Forager residential sites cause problems because they may be responding to different suites

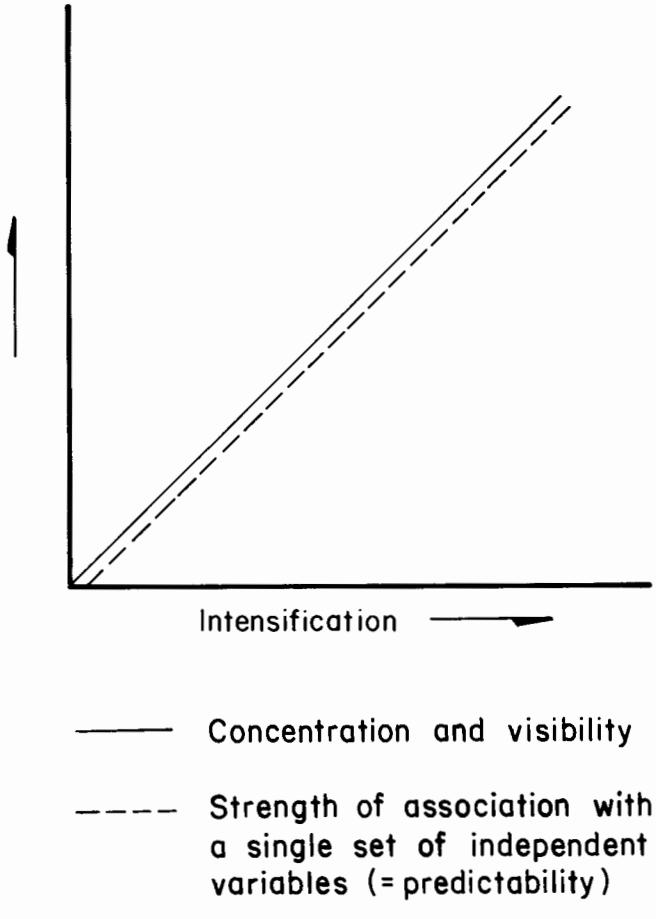


Figure 4.6. Suggested effects of increasing intensification on the concentration and, hence, visibility of archaeological materials at residential sites and locations within the foraging radius where nonexhaustible resources are exploited and on the strength of association of each of these site types with a single set of independent variables.

of environmental variables and because they fit the concept of "site" poorly. Residential bases in more intensified adaptations are less subject to these particular problems, but predictions about their locations also have complicating factors. Locations of these sites represent a response to an increasing number of variables, an increasing proportion of which (the locations of other contemporaneous sites, for example) cannot normally be used for prediction. This discussion makes clear the theoretical basis for Kohler and Parker's (1986) insistence on modeling different adaptation types in one area through time, but in order to do this we must be able to "sort out" the overlapping archaeological records.

Spatial Heterogeneity. Let us now briefly consider the effects of increasing spatial heterogeneity—patchiness—while ignoring intensification and temporal predictability. The aspects of spatial heterogeneity that have the most important implications for where sites will be located and how visible and predictable they will be are the degree to which the critical, nonsubstitutable resource patches overlap, the extent to which each resource type is concentrated, and the distance between patches of substitutable resources. First, we suggest that the strength of association between the distribution of archaeological materials and the distribution of a particular resource type (and therefore the *predictability* of those archaeological materials) should increase as resource patches

1. become more concentrated in space, so that equivalent resource-type patches are increasingly distant from one another; and

2. overlap more in space with other nonequivalent (nonsubstitutable) resource-type patches.

These proposed relationships are in accordance with common sense. The occurrence in a single location of more than one critical, nonsubstitutable resource (say fuel, large game, and roots) increases the likelihood of use, and reuse, for that location. If equivalent resource types (for example, carbohydrate resources with similar processing requirements and storage characteristics) are fairly continuous across the landscape, the strength of association between archaeological materials and any one of those resource types should be low. Equivalently, if patch size is very large, or if patches are close together, predictive success will tend to be low. It is important to remember that resources include things other than food; fuel is probably universally needed, but other "amenities," such as well-drained sediments deep enough to enable construction of a pithouse, may be peculiar to particular adaptations. We do not necessarily know the identity of these food or nonfood resources, however.

The visibility of archaeological materials, and to some extent the ease with which the concept of "site" may be applied, should increase under the same circumstances in which predictability increases. The same environmental circumstances that serve to bind environmental features and archaeological materials closely together should also serve to concentrate those materials into sites. It should be noted, however, that sites, even in these systems, are not the remains of discrete episodes of behavior. Because concentrated materials are easier to find than

dispersed materials (Wandsnider and Ebert 1984), concentration increases visibility of clusters. We do not advocate the use of the concept of "sites" without recourse to the entire explanatory modeling process and explicit recognition of the different meanings that the term *site* might have.

Where there is considerable spatial overlap among critical nonsubstitutable resources, a relatively small number of independent variables should adequately predict the presence or absence of archaeological materials. That is, where there is strong spatial correlation among the potentially important environmental variables, a few may successfully stand for many. Where spatial overlap among resources is low, a larger number of proxy environmental variables may be required for prediction. The general relationships suggested here between aspects of the spatial structure of critical resources and aspects of the predictive modeling process are graphically summarized in Figure 4.7.

Temporal Predictability. What, finally, are the effects of increasing constancy and contingency in the temporal distribution of various resources on the predictive process? Remembering that constancy and contingency can be summed to create a measure of temporal predictability, we propose that archaeological materials will be relatively concentrated, visible, and predictable in places where resources have either high constancy or high contingency; archaeological materials tend to be spatially predictable where resources are temporally predictable (Figure 4.8). (This prediction ignores concurrent variability in the spatial structure of resources; obviously, spatial concentration or dispersion of resources, as outlined above, also affects these relationships.) Places where both constancy and contingency in the temporal distribution of resources are low will not favor concentrated, repetitive, or long-term use and in general should not be associated with residential site types. High constancy of resource availability should favor low residential mobility, while high contingency should favor regular seasonal reuse. The coastal salt marsh/sea island/estuarine systems of Georgia and the Calusa area of southwest Florida are examples of environments with high constancy in the resources critical to human survival (Marrinan 1975). Most noncoastal North American environments experience greater seasonal pulses in temperature or precipitation, reducing the constancy of most critical biotic resources. The large rivers with their runs of anadromous fish and the root-gathering areas of the Columbia Plateau provide good examples of high-contingency environments.

The Interaction of Intensification, Spatial Heterogeneity, and Temporal Predictability

Finally, how do these three dimensions of variability—economic intensification, spatial heterogeneity, and temporal predictability—tend to interact? This is the important question for predictive modeling, since it is artificial to discuss these dimensions as if they were totally independent of one another. It seems obvious that certain kinds of spatial and temporal variability in resources require some intensification practices—particularly storage—before the resources can be exploited at all. Arctic adaptations to resources with low constancy, only moderate contingency,

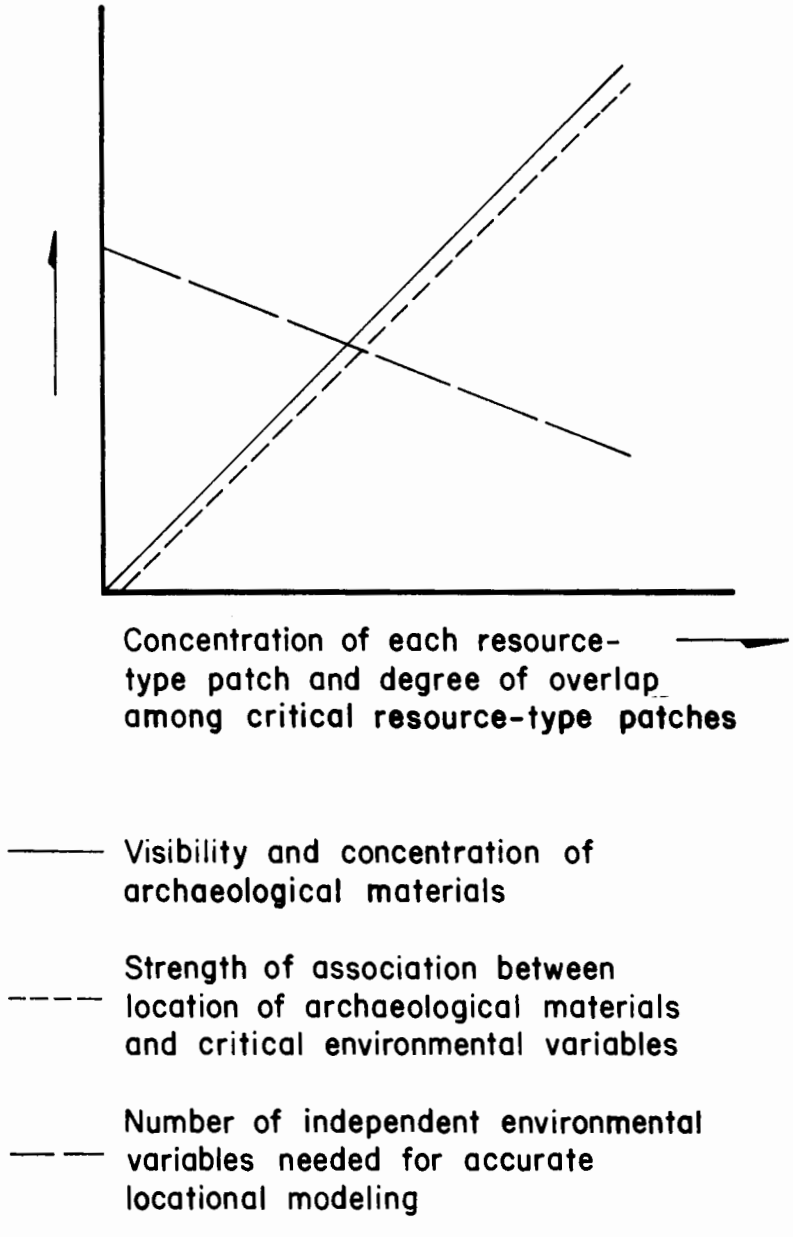


Figure 4.7. Suggested effects of the spatial characteristics of critical environmental resources on locational modeling.

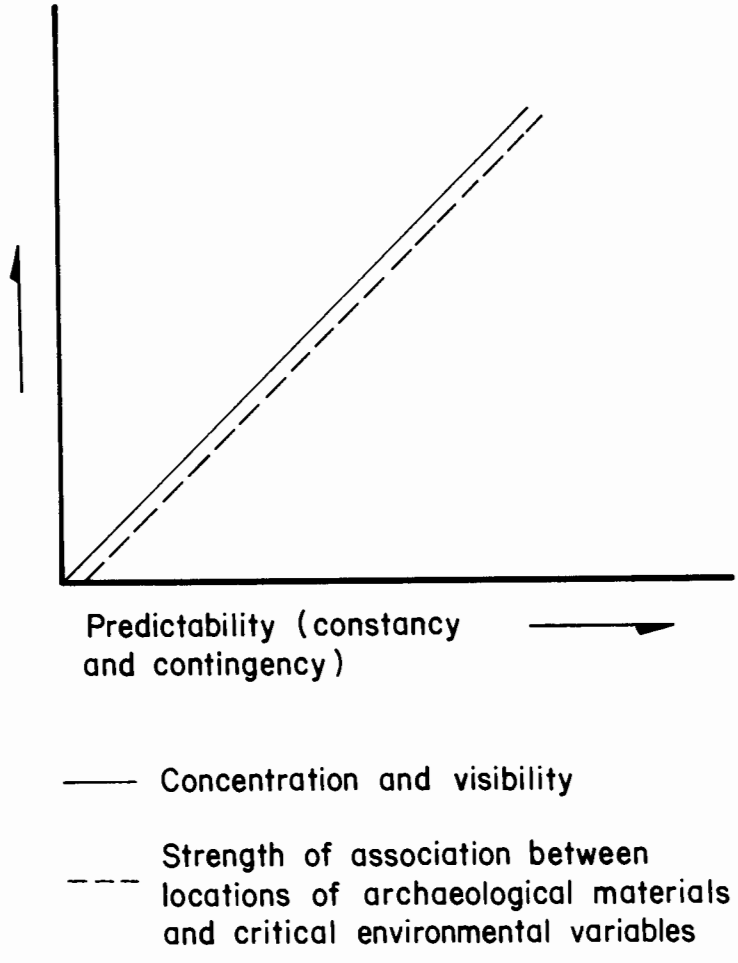


Figure 4.8. Suggested effects of the temporal characteristics of critical environmental resources on locations of archaeological materials.

and large distances between critical resources are good examples of this. Other combinations of environmental factors allow either a forager way of life or more intensified economies to thrive; under these conditions we might expect some historical tendency for the replacement of foragers by collectors and perhaps agriculturalists following the competitive exclusion principle (Bettinger and Baumhoff 1982; Kohler 1976). Relatively low constancy coupled with high resource productivity and relatively little spatial overlap in critical resources has seemed to favor intensification in many temperate portions of North America. This intensification involves increased population, increased packing, decreased residential mobility, increased storage, and even production of storable foods. Still other kinds of spatial and temporal variability discourage or select against intensification. Foragers in the tropical rainforest exploit resources that have high constancy and

high spatial overlap in the critical resources but low resource accessibility and little spatial variability. These environmental factors leading to high residential mobility, in conjunction with the prevalence of tropical diseases and pests, keep populations below the foraging capacity of these environments.

Implications for Inductive Empirical Models

We have suggested that various economic and ecosystemic factors affect *(a)* the number of relevant independent (determinant) environmental variables needed for accurate predictive locational modeling; *(b)* the extent to which factors in the natural environment, by themselves, will be adequate predictors of location; *(c)* the strength of the association between various predictor variables and the location of archaeological materials; and *(d)* the concentration and visibility of these materials. Let us now explore the implications of these suggestions for the inductive, empirical modeling of site location commonly practiced today.

First, there is no reason to believe that locations for all site types produced by all subsistence-settlement systems in all environments are equally predictable. Other things being equal, predictability (strength of association with critical environmental factors) should be relatively high in landscapes where equivalent resource-type patches are concentrated and isolated, have high overlap with other nonequivalent resource types, and have high temporal predictability. For residential site locations, accuracy of prediction (which is equivalent to the strength of association with relevant independent variables) should increase in more intensified economies. But the location of residential sites in such economies becomes an increasingly multivariate problem, and the independent variables affecting location increasingly include locations of other residential sites—information not typically available to or easily utilized by inferential predictive models.

Other problems for inferential predictive models involve the differential concentration and visibility of various site types in areas where the resources differ in spatial concentration, overlap, and temporal predictability, and in economies at differing levels of intensity. Residential bases become increasingly concentrated and visible under the same conditions that promote predictability, as reviewed above. Other site types may or may not become more visible under intensification, depending on their function and location in relation to a residential base. Other things being equal, we assume that site types other than residential bases will be underrepresented in samples from most modern and all older surveys.

Taking these points into consideration, it is unlikely that inferential predictive models will perform well in areas where resources are not concentrated, overlapping, and temporally predictable, or where residential sites have low visibility (such as those of foragers) or high locational dependence upon factors of the social environment. On the other hand, we can expect inferential predictive models to perform relatively well when the opposite conditions hold.

One point of this discussion is that we cannot expect inferential approaches to be equally successful in all circumstances, and even for all kinds of archaeological manifestations within a single settlement system. Nor is inferential prediction likely to be completely successful in most applications, both because of the complications outlined in this section and because of a certain indeterminancy in all living systems, including settlement systems. Another important point is that archaeologists need to begin to characterize the environments in which they work in terms of operationalized, consistent, ecosystemic factors, such as temporal constancy and contingency and degree of spatial concentration and overlap of resources, instead of by simple reference to the presence or absence of particular resources at particular points on the landscape. Admittedly, this will be difficult for even modern environments, let alone for paleoenvironments that differ from those of today, but we hope that this section has pointed out the necessity for such characterizations in understanding how settlement systems are structured and, therefore, how their positioning on the landscape might be predicted.

DISTRIBUTIONAL ARCHAEOLOGY

Approaches to Congruence Between Theory and Method

So far in this chapter we have discussed the effects on the archaeological record of differences in the organization of human systems, of a number of depositional and postdepositional processes, and of general ecosystemic (rather than single environmental) variables. We have tried to show the implications of these different determinants of the archaeological record for modeling and prediction. Some forms of organization and some temporal and spatial attributes of ecosystems lead to the formation of an archaeological record that is relatively more visible and predictable than records formed under other organizational and ecosystemic principles. We have suggested that the least visible and least predictable archaeological record is created by foraging activities—either foraging components of generally logistically organized systems or human systems whose subsistence activities are wholly organized around this mobility/settlement strategy.

What this means is that an expectably large proportion of the archaeological record left anywhere by all past peoples will consist of relatively continuous, low-density, low-visibility remains. Such an archaeological record cannot be dealt with using site-centered discovery and measurement methods; in fact, it may not even be detectable via traditional survey. In addition, the clustered materials that result from intensive reuse of circumscribed places (the things we think of as sites) are superimposed on this more continuous, lower-density record. In order to sort them out, to distinguish occupational and functional episodes from one another, we must record artifacts and features as a continuous phenomenon.

If in fact at least part of the archaeological record is continuous, and the ethnographic ("theoretical") as well as methodological arguments presented

throughout this chapter support that it is, then the meaningfulness of predictive modeling based on a "site" vs "not site" concept is called into serious doubt. One example of why this might be the case will be considered briefly here. In certain chapters of this volume it is argued that it is not enough simply to show that the locations of sites are highly correlated with the locations of supposed independent environmental variables; one must also show that "nonsites" (by which the author means places that are *not* sites or areas that do not contain sites) are less strongly or are negatively correlated with these same variables in order to allow "prediction" using such a variable.

How are such "not-sites" to be found? Unfortunately, the author continues, it is too expensive to look for them, but fortunately, according to him, we don't have to. Archaeological sites, he contends, are rare phenomena that only occur "about 1 percent of the time." Therefore, if one randomly chooses points at places where sites haven't been located through actual survey, it is to be expected that only 1 out of 100 points will actually be sites by chance, and the rest will be "not-sites." This argument is sometimes broadened further: in one geographic information system study Ebert knows of, the randomly selected "not-site" sample consists of *areas* 2 mi on a side, only 1 percent of which are supposed to contain sites by chance.

But just where would one have had to undertake a survey in order to think that sites only occur 1 percent of the time? Some people reply to this point by admitting, "Yes, you'll find archaeological materials everywhere you look, but not necessarily sites." And this is the real point: *How are sites to be distinguished from isolated occurrences or nonsites or not-sites?* By assuming we know that they are only "really" sites 1 percent of the time? By using different (explicit or implicit) definitions of sites vs whatever else in each survey, or even within a single survey?

Elsewhere Ebert (1986) has argued, at length, that one of the biggest problems that archaeology, particularly cultural resource management-directed archaeology, has is reliance on an unworkable, insupportable "site" concept. There are, thankfully, theoretically as well as practically valid alternatives to "site" approaches. These approaches are soundly based in archaeological literature and practice and are drawing increasing interest from both the archaeological and managerial communities. We would like, therefore, to conclude this chapter by offering an example of a methodological direction designed to record the continuous archaeological record. We believe that many such methodological innovations, critically informed by both general and middle-range theoretical concepts, will be needed before we can learn to predict characteristics of the archaeological record and locations of cultural resources accurately.

Background: Nonsite and Off-Site Archaeology

Recognition of the complexities of the formation of the archaeological record coupled with dissatisfaction with most traditional means of recording this record has led a number of archaeologists working in different parts of the world and with

different problem orientations to recommend new ways of approaching the spatial patterning of surface assemblages. One of the most promising of these involves what amounts to a reconsideration of the basic unit of archaeological analysis, which until this time has explicitly or implicitly been the site. David Hurst Thomas was one of the first to express dissatisfaction with the site concept in the literature, calling for a "nonsite sampling" (1975:61) approach. Certain sorts of depositional situations and certain problem orientations, he argued, make a site sampling approach not only "inessential, but even slightly irrelevant" (1975:62), and he suggested an alternative survey method in which individual artifacts, features, and other cultural items form the minimal operational units. This approach, used during Thomas's Reese River Ecological Project, was designed to test archaeologically the consequences of Julian Steward's model of ethnographic settlement patterns of the Great Basin Shoshoneans (Steward 1938). If Steward's model could be shown to describe the prehistoric case accurately, Thomas reasoned, then the contention of some anthropologists that historically observed Shoshonean behavior was due to acculturation in the wake of European contact would be disproved.

In order to study the ways in which "members of a single hunter-gatherer society moved themselves across the landscape, in a stable yet flexible pattern of transhumance" (1975:64), Thomas compared the cultural debris left by these people in each of a number of "microenvironments" or sampling strata in the Reese River Valley in Nevada. Locations and characteristics of individual artifacts were recorded, and artifact-density statistics were used to analyze some aspects of the prehistoric systems represented. Although the relationship between these observations and the human behavior that created the data was not explicitly defined, Thomas's work remains a provocative illustration of methods of data collection and analysis that are not totally dependent on the site as an analytical unit.

Bettinger (1977a, 1977b) employed Thomas's methods of density analysis in a similar inquiry into the correspondence between ethnohistorically observed behavior and the patterning of surface assemblages in eastern California's Inyo and Mono valleys. Although both Thomas and Bettinger advanced sound theoretical reasons for their nonsite approaches, it is likely that the nature of the observed archaeological surface record in their study areas was more than a little responsible for shaping their research designs. In much of the arid and semiarid American West, surface archaeological remains consist of large expanses of sparsely distributed artifacts and features that can only be sorted into discrete sites by means of arbitrary boundary-setting criteria.

Another environment that is archaeologically similar to the American West is the arid belt of East Africa extending southward from Egypt through the Rift Valley. At approximately the same time that Thomas and Bettinger were working in the Great Basin, archaeologists in Africa were beginning to develop their own methods of measuring diffuse artifact distributions. Faced with the sparse and probably disturbed artifactual evidence from the Acheulean in Kenya and Tanzania, Glynn Isaac and his colleagues approached the archaeological record from a consideration of natural depositional and preservation processes (Bunn et al. 1980; Isaac

1966, 1967, 1978; Isaac and Harris 1978). These studies were oriented toward assessing the patterning of artifacts within sites and on occupation floors believed at the time to be the consequence of single behavioral episodes. Isaac was also concerned with nonsite distributions of cultural items, however—the "scatter between the patches" (Isaac and Harris 1975) that makes up a large proportion of the total number of cultural items discovered in large areas of arid East Africa.

As with the American nonsite strategies, sample quadrats were defined along the eastern shore of Lake Turkana, and the locations of individual artifacts and features within these sample areas were recorded. Isaac (1981) argues that the density and patterning of various artifact types are related to prehistoric mobility patterns analogous with those of present-day hunter-gatherers. Going a step further toward the reconciliation of nonsite and site-oriented archaeology, Isaac and his colleagues have more recently suggested what must be seen as yet another alternative unit of analysis, the "mini-site" (Isaac et al. 1981:105). Although the term may be unfortunate, the implication that the remains of many past behavioral events or series of events might consist of very small or diffuse assemblages is worthy of consideration.

Perhaps the most systematically developed approach to understanding the meaning of archaeological surface assemblages employing the artifact as an analytical unit is Robert Foley's "off-site archaeology" (Foley 1980:39–40). This methodology was the result of Foley's attempts to compare site locations and the distributions of resources in a catchment area or "home range" around a site (Foley 1977). Starting with the assumption that resource usage is distance dependent, Foley proposed a model in which a study area would be gridded into squares and the total relative resource productivity for each area would be calculated on the basis of detailed ecological field studies. Next, given the location of a site of interest, isocals or areas with consistent extractive values for that site would be drawn. All those areas in which the availability/cost ratio for resources was positive would be considered to be likely candidates for the home range for that site (Foley 1977:178).

Operationalizing such an explicit economic model would, of course, require a detailed knowledge not only of all relevant prehistoric ecological parameters but also of the locations in space of all sites or localities participating in the cultural system of interest. In interpreting the preliminary results of archaeological survey undertaken in the Amboseli Basin in Kenya, Foley recognized that artifacts seemed to be "distributed ubiquitously across the landscape. In contrast to this, demonstrable primary stratified sites are extremely rare" (1980:39). This was due, he felt, to at least two broad classes of processes: those arising from the patterning of actual human behavior in the past, and those created by postdiscard taphonomic, depositional, and postdepositional forces working upon the discarded artifacts. Later consideration of the formation stages for the archaeological record led Foley to draw a number of inferences upon which the necessity for and methodology of off-site archaeology were to be based (Foley 1981c:31):

1. Sites are nodes in a continuous distribution of archaeological materials.

2. Home-range behavior provides the theoretical underpinning for continuous archaeological regional distributions.

3. The processes of continued occupancy, leading to accumulation of materials, and postdepositional mechanics compound the continuous distribution, as well as increasing its complexity.

4. The artifact, the basic unit of an archaeological distribution, can and should be used as the unit of regional analysis.

The methodology employed by Foley in his Amboseli survey was developed to ensure the collection of data—artifact locations and the characteristics of these artifacts—on items occurring continuously but not uniformly across the landscape. The study area was sampled using methods tested by plant ecologists who have "the same problem of integrating small analytical units or data objects (plants, artefacts) with large survey areas" (Foley 1981c:34). The sample areas encompassed 0.05 percent of the total study area.

Two basic classes of data were collected in the sample units. First, the natural and particularly the preservational and depositional environments were recorded. Sediments were classified, and the natural processes acting on them (erosion, compaction, topographic effects, vegetation cover, and animal or recent human activity) were noted. Next, artifacts were recorded in terms of raw material, size, artifact or flake type, platform, cortex, and condition; taxonomies for pottery and associated bone were also devised. In addition to the surface survey, a number of experiments designed to test the short-term effects of rainfall, erosion, compaction, and other taphonomic processes were undertaken. The exact locations of artifacts within the sample units were apparently not recorded—a very significant omission that, coupled with small sample unit size (5 by 50 m), precludes any but the grossest density-based spatial analyses.

Data on the occurrence of artifacts in the sample units were extrapolated to the entire study area, and density contours were drawn. Other contour maps also extrapolated from the sampled areas to the total study area depicted densities of raw materials and artifact types, proportion of cores to other artifact types, artifact length and width, occurrence of retouch and edge damage, and other artifact characteristics. Foley's analysis of his Amboseli data, like his earlier work (1977), proceeded from a goal of examining humanly important aspects of the environment. He attempts to do this by formulating models that predict the areas of most intense use by past groups with pastoralist and hunting-gathering adaptive strategies and then testing these predictions using artifact density data.

These pioneering efforts to arrive at congruence between theoretical ideas about the formation of the archaeological record and methods of discovering, measuring, and analyzing cultural resources inspired two recent experiments with adapting nonsite or "distributional" archaeological survey to cultural resource management. These experiments are discussed below.

Distributional Archaeology: Paths Toward Theoretical/Methodological Congruence

In order to apply the theory-based explanatory framework examined earlier in this chapter to the archaeological record, it is absolutely essential that there be a congruence between theory and method. Such approaches as nonsite or off-site archaeology, in which the artifact is the unit of discovery and analysis, are certainly a step in the right direction when we are attempting to deal with continuous aspects of the archaeological record, as discussed above. There are shortcomings in these approaches, however. One of the chief problems with recent artifact-oriented approaches has to do with sampling. If it is the patterning in the continuous distribution of archaeological materials that we must measure, then the best way to do this is by choosing a relatively large "window" through which to look—by surveying for, discovering, measuring, analyzing, and interpreting archaeological materials over relatively large, contiguous sample units.

The remainder of this section will describe what one of the authors and his colleagues (Ebert et al. 1983) have referred to as *distributional archaeology*. Distributional archaeology is a nonsite-oriented approach that yields data that are congruent with the theoretical concepts of mobility and artifact discard presented above. Distributional archaeology has been carried out in two different governmental contexts as this volume goes to press. In 1983 the Bureau of Reclamation and the National Park Service funded a distributional survey at and around Fontenelle Reservoir in southwestern Wyoming, and the Bureau of Land Management Las Cruces (New Mexico) District recently conducted a distributional survey near El Paso in conjunction with the Navajo-Hopi Land Exchange.

Unfortunately, no detailed accounts of these surveys have yet been published, although a number of papers and reports are available (Ebert 1983a; Ebert et al. 1983; Larralde 1984; Wandsnider and Ebert 1983, 1984; Wandsnider and Larralde 1984). These papers have been compiled in a report edited by Drager and Ireland (1986). This section will not provide an exhaustive discussion of this methodology but rather will summarize some of the main points. Distributional archaeology is by no means fully perfected, and experimentation with similar approaches should be encouraged.

Distributional archaeology was conceived with several major objectives in mind. It is oriented toward the relatively complete and continuous survey of archaeological materials—artifacts and features—over large contiguous areas. Large areas relative to the scales of the archaeological patterning must be surveyed, and their contents analyzed, if we hope to sort out overlapping distributions in the continuous archaeological record. The distributional archaeology methodology calls for discovery of artifacts and features through intensive surface survey, recording of the location of each artifact or feature as a point in space, and consistent in-field coding of artifact attributes. All artifacts, including nondiagnostic tools and debitage, are recorded.

The Seedskadee Project

Survey Design

The Bureau of Reclamation/National Park Service survey around the Fontenelle Reservoir was called the Seedskadee Project (Figure 4.9). This survey was designed as an experiment to test systematic survey methods for recording the continuous archaeological record. The survey design was directed by two major propositions: *(a)* units of analysis and discovery structure the ways in which archaeologists think about the nature of the archaeological record and, in fact, what is found during fieldwork (Binford and Sabloff 1982); and *(b)* very little is known about what the archaeological record means or what it looks like. For these reasons, the units of analysis employed during the survey had to be units with little or no meaning already attached. Individual artifacts were chosen as the units of discovery and mapping; attributes of artifacts were chosen as the units of data recording. The discovery and recording methods used were carefully designed to minimize biases in what was recognized as an artifact, what data were considered to be appropriate to record, and how those data were recorded.

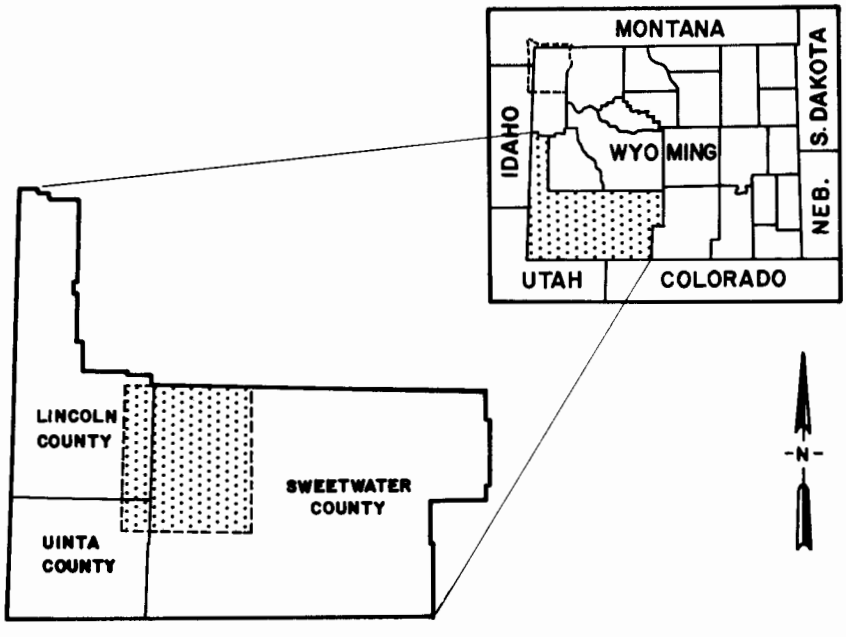


Figure 4.9. Location of the Seedskadee Project, a distributional (nonsite) archaeological survey undertaken by the National Park Service and the Bureau of Reclamation in southwestern Wyoming.

A simple random sample of 25 500 by 500 m sample units was surveyed during the Seedskadee Project (Ebert 1983a). The sample was not stratified by environmental zones because the zonation present was thought to represent differential surface geomorphological processes rather than past natural conditions. Responsibility for data recovery was delegated to three separate crews. A five-member discovery crew was responsible for finding and flagging artifacts and for maintaining even ground coverage in precisely controlled 5 m transect intervals. The data-recording crew consisted of three individuals who numbered the pinflags marking artifacts and recorded artifact attribute data in a format designed for easy computer input after the fieldwork phase of the project was completed. The two-person mapping crew was responsible for provenience control of artifacts, most of which were mapped individually using an electronic distance measuring (EDM) device and a prism. In areas where artifact density was very high, mapping of individual items was abandoned, and 1 m grids became the provenience unit.

When additional artifacts were found by the recording crew, they were flagged separately. The distributions of these later finds often resembled the results of traditional site surveys in that they tended to be far more clustered than the distributions marked by the discovery crew. As a rule, highly visible artifact concentrations received more attention than interlying areas, as is the case with traditional survey methods. The items found by the recording crew often doubled or tripled the number of artifacts recorded in a sample unit.

General Results of the Seedskadee Distributional Survey

The end product of these survey procedures is a data base that consists of some 170,000 coded attributes, predominately locational data and lithics descriptors from 17,000 artifacts. Analysis of the Seedskadee data base, emphasizing the search for spatial patterns among attributes, is presently proceeding along lines that will be discussed below. Some preliminary impressions gained from the Seedskadee experiment, however, have immediate implications concerning the appropriateness of the approach and the nature of the contributions that it can make to predictive attempts:

1. There were prehistoric artifacts in all environmental zones. They occurred in differing (but usually unexpectedly high) densities and in many different kinds of distributions that appear to vary in both spatial configuration and content. It seems that the kinds of distributions encountered at Seedskadee would confound the usual methods of doing predictive modeling (i.e., defining environmental parameters for site location) because the data base is gradational in distribution and density, rather than made up of discretely bounded "sites."

2. The harder one looks, the more one finds. Although this is a simple observation, its repercussions for management of archaeological resources are profound, since RFPs generally emphasize acres surveyed rather than cultural resources located per dollar spent. The perception that archaeologists have of the archaeological record is a direct function of the context of discovery: survey

interval; time spent on sweeps, on flagging concentrations, or on recording contents of grid squares; and external and internal crew goals and conditions.

It was also observed that *surface* and *subsurface* are relative, dynamic terms. This point is easily illustrated in areas like dunes, where the acts of discovery, mapping, and data recording change the surface archaeology: artifacts are buried and uncovered through scuffling and trampling during the course of the survey itself. Noncollection survey is often (probably always) destructive of the archaeological record. Not only does the survey have a direct impact on the location of artifacts, it is likely that indirect impacts, such as alteration of the soil's surface and of vegetation, will affect the rates and nature of local natural processes in the future.

3. Error, variability, and sources of bias in method and results must be evaluated and explained. To address such problems, two control experiments were included in the project to help in the evaluation of data reliability. In the first, a sample unit was seeded with "pseudofacts": nails and washers painted to approximate the color of the ground and natural lithic materials occurring in the area. Some of these items were distributed in clusters or "sites," while others were placed individually as "isolated occurrences." These were flagged and recorded by the discovery crew, which yielded information about accuracy of the discovery procedures. Approximately 55 percent of the pseudofacts were recovered by the discovery crew at a 5 m transect spacing, with an additional 10 percent being found by the follow-up analysis crew. More interesting, however, were the proportions of clustered vs isolated pseudofacts found. The discovery crew located 68 percent of the clustered artifacts but only 16 percent of the isolated items (for the analysis crew the figures were 12 and 6 percent, respectively; Wandsnider and Ebert 1984).

In a second methodological experiment, a purposefully manufactured lithic assemblage was independently coded by the three principal data recorders. There was considerable inconsistency among coders even though they inspected the assemblage at the same time under the same conditions. It is possible to control for such inconsistencies if their extent is known, however, and procedures for doing so are discussed at length by Larralde (1984).

4. With a systematically organized, multicomponent survey team such as the three-part Seedskadee crew, portions of the crew can complete their individual tasks at their own speed and under ideal conditions, and this greatly increases the yield of actual product (in terms of information) per person-hour worked. In a period of approximately seven 10-person weeks, some 170,000 attributes were recorded. This is the information equivalent of 2–3000 of the most detailed site recording forms in use in the United States. Although the amount of ground covered during this time (625 ha or 1544.35 acres) is less than for most traditional, site-oriented surveys, the information yield is high. The information-yield argument is very important when considering the cost-effectiveness of any in-field data-collection program.

The question might be asked, of course, just what the real "information equivalence" is between 170,000 artifact attributes and the data contained on 2–3000

detailed site forms. In order to answer this, what those data consist of must be explicitly considered. Distributional archaeological data consist of known point-locations and characteristics of the physical materials that make up the archaeological record. Site data in almost all cases consist of hastily formed opinions about abstract boundaries of supposed past occupations, guesses as to how many artifacts might be found within those boundaries if one were to count them, a list of diagnostic materials found during a walk around the "site," and the surveyors' enumeration of the cultures that occupied the area in the past and what the members of those cultures were doing there (camping, chipping stone, hunting, etc.). We would suggest that in most cases "information equivalence" isn't even the right framework for such a discussion. The difference is information vs abstractions.

5. Even though finding sites is not the point of a distributional survey, the results of spatial clustering routines run on the Seedskadee data suggest that the distributional survey discovered more "sites" than recent traditional surveys in the immediate project area. This is true even if allowance is made for the intensity of survey. The Seedskadee survey was 3–6 times as intensive as 15–30 m transect interval surveys done recently in the area (Reynolds 1983); our first impression is that the Seedskadee distributional survey located from 10 to more than 50 times as many sites as the traditional surveys did. This means either that linear or sinusoidal intensity-to-yield models of surface survey results such as that presented by Judge (1981) are unwarranted or that we did not reach the hypothetical falloff point even at a 5 m transect interval. Are even smaller transect intervals necessary in certain situations?

6. Field observation during the Seedskadee Project revealed that the scale of patterning of the natural processes that affect the visibility, preservation, and integrity of the archaeological record are of a very local nature. These processes are controlled by local topography and other small-scale factors and thus are often smaller in scale than culturally caused clusters of artifacts. As discussed above, it is necessary to factor out the effects of natural depositional and postdepositional processes before one can decide what cultural patterning looks like. This means that extremely localized, small-scale geomorphological mapping and process measurements over time may be absolutely necessary before any predictive modeling of artifact or site distributions can be done.

The Navajo-Hopi Land Exchange Project

Another example of a distributional archaeological survey in which the site is not the explicit unit of either recording or analysis is the Navajo-Hopi Land Exchange Project survey, conducted by the Bureau of Land Management just west of El Paso. This survey is much larger in scale than the Seedskadee Project and represents several refinements on the methods used in the earlier survey.

The Navajo-Hopi Land Exchange (NHLE) survey was conducted in three adjacent survey areas, which together comprise some 16,000 acres (25 mi^2) of mesa top and breaks. Much of the area is covered by a thin sand mantle exhibiting coppice dunes and blowouts. Previous site-oriented research in the study area recorded sites with adobe pueblo structures as well as many scatters of lithic and ceramic materials and isolates. The adobe pueblos are typically not visible unless they have been disturbed by natural or cultural processes, although their associated middens have high artifact densities and thus high visibility.

Phase I of the NHLE survey was designed primarily to fulfill management needs. Its goal was to determine, through a relatively low intensity transect survey, which areas contain dense concentrations of resources, particularly structure-associated resources, and should therefore be excluded from the land exchange and preserved. Phase I was also expected to identify areas in which the nature of the archaeological remains did not warrant automatic exclusion but did require further survey, study, and possible excavation prior to the land exchange.

During Phase I, 400 by 400 m and 800 by 800 m sample units (totaling more than 60 km^2) were surveyed at 25 m and 50 m transect intervals. All artifacts and features occurring within 1 m on each side of the surveyors were tallied for each transect, and densities of materials were calculated along each transect. These density data were analyzed using a clustering technique in which areas were examined on the basis of whether they contained portable and/or nonportable containers, portable and/or nonportable implements, and low- and/or high-volume processing facilities. Predictions were then made as to which areas should contain structural remains. An independent, structure-oriented discovery survey was carried out, and the Phase I density analysis was found to have been very successful at predicting which areas would contain subsurface structures.

Two other classes of areas were also isolated during the Phase I survey: those with very low densities of cultural resources and those with moderate densities of artifacts and features but without associated structural remains. These areas are the subject of Phase II, an intensive survey similar to that described for the Seedskadee Project (Camilli et al. 1988). In this phase 13 400 by 400 m units and five 800 by 800 m units were studied using a 5 m transect interval. Individual cultural items (artifacts, features) were the unit of discovery, mapping, and analysis.

Certain cultural resources, including unifaces, bifaces, and rimsherds, were collected during this phase, and some of the areas with surface features, such as firecracked rock and hearths, and some scatters with no features were excavated. Generally, however, artifact and feature analysis carried out during the course of Phase II was done in the field. The Phase II in-field coding taxonomy was directed toward not only the identification of formal tools or diagnostic materials, but especially toward identification of lithic production strategies.

Artifact Coding and Analysis

We need to touch briefly upon a very important subject—what data need to be coded when artifacts are used as the units of discovery and analysis. It is often suggested that all archaeological research takes place in unique situations and that each researcher's problems are different, and therefore that no hard-and-fast rules can be formulated as to the field methods and analyses archaeologists should use. It may be that any area containing cultural resources is unique on a very specific level—just as the distribution of molecules in Maxwell's glass of water was unique. We would suggest, however, that it is not the unique aspects of the archaeological record that are of interest, but rather those aspects that can be compared and contrasted from place to place—the *general* attributes of archaeological materials.

The practice of separating assemblages on the basis of formal attributes of diagnostic artifacts and labeling each of these as a different culture type or tradition defeats any attempt to recognize the differentiated portions of human organizational systems and thus precludes successful explanation, modeling, and prediction. Methods must be found for recognizing different systemic components and their overlap. Although some possible directions for this will be discussed below, we do not, unfortunately, know at present which general attributes of archaeological items are important in explanation.

It is possible, however, to describe a general direction that might be followed in determining how to code attributes of archaeological materials. A human system is composed of, among other things, a series of places where things are done. The key word here is *series*, and this chapter began with a discussion of the ways in which events at each place in the series are important to the operation of the entire system. Another set of components in a human system is technological items, which are also used in a serial way. Items used at places are sometimes discarded and at other times are modified there and used for other functions. Still other times, items are curated and taken away to be used at one or probably more different locations. Attributes that provide possible clues to the serial nature of technological strategies are, then, of major importance in understanding the components of systems. Such attributes include not only formal tool designations but also data on the nature of what most researchers class as debitage—utilized and unutilized parts of tools, and debris from lithic reduction, modification, and manufacturing.

Analyzing Data from Distributional Archaeological Surveys

It is necessary to establish linkages between the archaeological record and the organization of the past human systems that created this record before we can make successful predictions about the locations of cultural resources or about their meaning, usefulness, or significance in archaeological terms. Previous sections of this chapter have worked downward through the explanatory framework of archaeology presented in Figure 4.1, beginning with higher-level, theoretical ideas

about the nature of human settlement/mobility systems and middle-range theoretical ideas about technological strategies, discard behavior, and the natural formation processes that also affect the archaeological record. In this section, archaeological method has been considered and suggestions have been made about ways to discover and measure the archaeological record that are congruent with this theoretical framework—nonsite or distributional surveys yielding high-resolution spatial and attribute data collected from artifacts and features, rather than from sites. Given such data, however, what should we do with them?

This is not an easy question to answer, for it is one of the areas in which the most concentrated archaeological research has yet to be done. During the past decade, however, there has been some experimentation in the spatial and content analysis of assemblages, primarily intrasite analyses attempting to isolate activity areas within sites. Intrasite analysis is not *exactly* the same thing as what we will ultimately want to do with distributional archaeological data, although the intrasite analysis literature should suggest some ways in which archaeological data must be analyzed before we can understand patterning in the archaeological record produced by the action of past systems.

Carr (1981, 1984) suggests that prior to intrasite assemblage analysis it is necessary to differentiate carefully between activity sets (archaeological materials used together in space and time in the past) and depositional sets (those materials that aggregate in the archaeological record), since disjunctions between these two entities result in clusters of tools or implements that are not automatically equivalent to activity areas. Associations in the archaeological record may be the result of implements having been used together, but they can just as well be a result of overlap of activities through time or of natural depositional and postdepositional processes that cause polythetic, overlapping depositional sets (Carr 1984:120). The archaeologist faced with comprehending the intrasite archaeological record must, according to Carr, use these depositional sets to define *(a)* the spatial limits of activity areas and *(b)* the organization of artifact types into tool kits. These have been the goals of a number of archaeologists using various methods of intrasite analysis.

Wandsnider and Larralde (1984) break down contemporary intrasite archaeological assemblage analysis methods into three basic types. The first of these was developed by Robert Whallon at the University of Michigan, who became one of the pioneers of intrasite spatial analysis with the development of his dimensional analysis of variance (Whallon 1973) and comparison of its results with nearest neighbor analysis (Whallon 1974). Whallon's more recent work (1984) uses a more comprehensive spatial method called "unconstrained clustering." Unconstrained clustering identifies areas within sites that have similar assemblages by *(a)* constructing density maps for each artifact type, *(b)* calculating the relative proportion that each artifact type contributes to the assemblage at points across the site, *(c)* identifying similar assemblage types, *(d)* mapping the cluster members and examining their distribution, and *(e)* reconstructing the activities that occurred on the site in light of spatial patterns identified ethnoarchaeologically. Carr (1984) has criti-

cized Whallon's unconstrained clustering method because it assumes that activity sets are the result of single episodes of functionally similar behavior (i.e., they are monothetic [Carr 1984:136–137]), that tools are always discarded expediently, that no disturbing postdepositional processes occur, and that activity areas do not overlap but have sharp borders. These assumptions are probably unfounded in most if not all cases of human behavior and archaeological record formation.

Carr (1981, 1984) proposes new techniques that he feels overcome some of the problems with Whallon's and other spatial analysis approaches; these techniques describe the distribution of each artifact type within the site. Although Carr's process is too complex to be described in detail here (see Carr 1981), he uses point distributions rather than grid cell counts and employs digital filtering, Fourier analysis, spectral analysis, and histogram equalization, techniques in common use in the processing of imaged remote sensor data. Such technological means may hold great promise for archaeological pattern recognition.

A third class of approaches to intrasite spatial organization is exemplified by the work of Kintigh and Ammerman (1982) and Simek and Larick (1983). Kintigh and Ammerman's heuristic approach to spatial analysis combines "the sophistication of intuitive approaches with the information processing capacity and systematic benefits of quantitative treatments" (1982:31). This method divides artifacts into types and subjects the distributions of each type across space to a *k*-means nonhierarchical divisive cluster analysis. The archaeologist, using internalized knowledge about the scales and nature of archaeological formation processes, decides intuitively upon a cutoff point for the number of clusters of each type formed and then recombines the clusters of different types into a series of overlapping clusters that presumably represent activity areas.

While all three approaches to intrasite spatial analysis hold promise, they are all also directed toward specific reconstruction of the things that went on within sites. Before these methods can be applied to the continuous archaeological record across landscapes, the scale of application must be increased far beyond that discussed by these authors. Wandsnider and Larralde (1984) have also pointed out that each of these three approaches solves only some of the problems of spatial analysis: Kintigh and Ammerman and Simek and Larick only identify spatial clusters; Carr describes and compares the spatial organization of artifacts; and Whallon describes and compares assemblage content. Wandsnider and Larralde call for methods that permit the description and comparison of assemblages both in terms of content and in terms of spatial organization or structure, and they suggest a five-part method building upon archaeological theory as well as inductive statistical procedures:

1. *Development of an artifact taxonomy.* This could proceed in several ways: along deductive lines, based on ethnographic and ethnoarchaeological information; on the basis of experiments in lithic manufacture that identify the stages of artifact production and sequential use; on the basis of information about the mechanics of artifact function (edge angles, etc.); or through purely statistical and inductive clustering algorithms.

2. *The identification of spatial units.* Directed at defining the boundaries of assemblages, this stage of analysis might be based on interactive, heuristic techniques, such as those suggested by Kintigh and Ammerman (1982) and Simek and Larick (1983), which incorporate not only behavioral knowledge but also information on depositional and postdepositional processes. Perhaps the best way to define small-scale patterning in natural formation processes is through the use of remote sensor data, since these data could be machine processed at the same time as distributional information on clusters.

Other sources of information to be incorporated at this stage may come from a consideration of the characteristic shapes and sizes of the spatial patterns of human behavior. During fieldwork with the Nunamiut Eskimo, Binford (1978) identified a number of different zones of activities and artifact discard, including drop zones, toss zones, hearth-centered activity areas, and structure or tent scatters. Recognition of these patterns within overlapping distributions might be accomplished mechanically by varying grid frame sizes during analysis or by constructing shape-recognition filters. The larger zone types might be appropriate for discerning the boundaries of assemblages.

3. *Content description and analysis.* Once assemblages have been defined, their contents might be described on the basis of the taxonomy or taxonomies devised in stage 1, by means of such techniques as principal components analysis, which is used to compare the composition of different assemblages. Recently, Kohler and Blinman (1987) have proposed using multiple linear regression to estimate the absolute and relative contributions of several different periods of use and deposition to the total archaeological deposits at ceramic-containing sites in the Dolores River Valley in Colorado. Their technique is similar to Stahle and Dunn's (1982) use of multiple linear regression to estimate the contributions of various stages of bifacial reduction in a mixed collection. Both of these applications are aspatial, but they contribute to an understanding of the composition of mixed collections in terms of predefined constituents and might be used to sort out overlapping activity sets.

4. *Structural description and analysis.* While stage 3 is directed toward describing the contents of assemblages, stage 4 provides a description of the spatial organization within assemblages. This is where the smaller-scale "zones" characteristic of human activity might be recognized within the overall assemblage composition through digital filtering. It might also be possible to use small, simple clusters of materials that seem to result from single discrete activity episodes to design "filters" to pass through larger, denser, and probably more composite artifact distributions. Smaller, single-occupation clusters might be expected to exhibit more central distributional tendencies and higher correlations between artifact types in space than the larger, more composite distributions. Other filters might consist of sample frames of varying size that could be passed through complex distributions in the manner of Whallon's dimensional analysis of variance (1973, 1974). Wandsnider and Larralde (1984) also suggest that the spatial organization of the different principal components might be inspected.

5. *Pattern dissection* would constitute the last stage in distributional spatial analysis, according to Wandsnider and Larralde (1984). Larger and more complex assemblages (i.e., things found together—depositional sets in Carr's terminology) are undoubtedly the result of the complete or partial overlapping of many behavioral episodes. It may be possible to separate these episodes from one another, and certainly this is a necessary step in comprehending the complex systemic mechanisms that resulted in the archaeological record at any place.

It is quite possible that some of the procedures suggested by Wandsnider and Larralde (1984) might be implemented in different orders or as combined steps rather than separately. Some of them may also be unnecessary—for instance, stage 2, in which the boundaries of assemblages are sought. We may never really see bounded assemblages in the continuous, overlapping archaeological record but rather may be looking at portions of these through the "windows" provided by our sample units, by our survey area boundaries, or by natural surface processes.

The Solution: Dedicated Research Using Distributional Data

It is clear from the foregoing that two general things can be said for archaeological spatial analysis. The first is that archaeologists do not quite know how to do it yet, at least in ways that are congruent with the higher-level and middle-range theoretical ideas that we have about the formation processes of the archaeological record. The second is that spatial analyses directed toward understanding the complex, composite archaeological record will probably combine modern techniques such as digital image processing—some of which are just now being developed to the point that they will be useful to archaeology—and deductive reasoning in a complex interactive process. This process will draw upon both archaeological and ecosystemic theory to arrive at successful archaeological explanation and thus prediction. Such archaeological analysis is presently a goal rather than reality, a goal toward which both management and archaeological interests should be energetically directed.

SUMMARY

This chapter has been concerned with the method and theory of using anthropological explanation to predict things about the organization of past human systems as well as about the archaeological record. The explanatory process illustrated in Figure 4.1 involves the advancing of *models* that are used as the basis of prediction. While at first it might seem overblown to introduce anthropological explanation into a discussion of "practical" archaeological prediction, it has been argued and illustrated here that it is only in the context of explanation and explanatory modeling that archaeologists and managers can hope to make truly successful predictions of the locations and other characteristics of the materials that

make up the archaeological record. This is so because the things that determine the locations of the materials that make up the archaeological record are not static, unchanging properties of the environment that can be measured easily from topographic or environmental maps.

The archaeological record is not the same thing, or even the same kind of thing, as the way that past individuals dealt with their environment and the locations at which they dropped artifacts. The data that archaeologists collect, analyze, and attach significance to are the product of long-term use of the landscape. Large numbers of people, organized in different ways, have serially located their activities across this landscape, manufacturing and differentially discarding artifacts in ways that changed as the landscapes changed with paleoclimatic flux, and as the mobility and technological strategies within their cultural systems changed.

This chapter has advanced a general model of human subsistence and mobility strategies that vary along a continuum of intensification from a generalist, *foraging* strategy through a specialized, *collecting* organization. This model is not intended to represent the "whole truth" about past systems. Nonetheless, it provides a basis for making predictions, and if these predictions prove to be consistent with the observations about the archaeological record, this would tend to support the usefulness of the model. If the predictions made on the basis of this model are not supported by observations of the archaeological record, then an alternative model or models should be devised. This may be one of the most important problems currently facing archaeologists today—to arrive at and attempt to confirm models concerning the operation of past systems. This task lends significance to the discovery and conservation of archaeological materials, and it is therefore the reason why cultural resources should be managed and preserved.

Before archaeological data can be called upon to support or negate any explanatory model, however, the archaeologist must take into account the things that alter or otherwise affect the ways that we see the materials that past human systems discarded. These factors are also illustrated in Figure 4.1 at the beginning of this chapter.

There are two basic types of things that happen to the objects that human systems culturally modify and then discard or abandon. The first of these lies in the realm of natural processes, which incorporate discarded materials into the earth's surface and subsurface deposits and which act to preserve, rearrange, or destroy these materials. Natural processes also make archaeological material visible to archaeologists and managers, so that we know they are there and need to be conserved and studied.

The other factor affecting archaeological materials is that they are discovered, measured, analyzed, and interpreted by archaeologists. This is the realm of archaeological methodology. It has been suggested in this chapter that, in order to be successful at discovering those things we need to know about the archaeological record in order to be able to predict its locations and characteristics (and thus its

significance in management terms), archaeological methods must be compatible with the theories we have about the ways in which materials were discarded by past human systems. It has also been argued that this is not presently, at least very often, the case, and that we may need to alter significantly the ways in which we deal with the archaeological record today as archaeologists and managers.

Another very important area of archaeological methodology concerns the natural phenomena that we measure and compare with the distribution of archaeological materials as they presently exist. These variables must be chosen in accordance with our ideas about the organization of past human systems if they are to be useful in predicting the characteristics of the archaeological record. The things that biologists, ecologists, and the people who make topographic maps have measured may not be the best variables to use if we wish to elucidate the organization of past systems; we have discussed the alternative of using ecosystemic variables in archaeological explanation rather than relying on specific resources, species, landforms, or other convenient proxy "indicators." In order to use ecosystemic variables in our modeling and predictions, we may have to do most of the measurement work ourselves.

Many archaeologists may disagree with the models of past systems organization that have been advanced in this chapter and with our suggestions about the relationships between these models and ecosystems variables and about the consequences of these relationships for the archaeological record. That is good, for it gives us all something to think about and to try to build upon and to alter so that it "fits" the archaeological record that we discover and deal with. There are few archaeologists, however, who will argue that we do not need to model past systems organization to predict the locations and nature of the archaeological record that we are all concerned with conserving.

This chapter, therefore, should not be thought of as advancing any particular model or models that will best typify what human systems were like in the past, or how they were related to the world in general. The theme of this chapter is instead that it will *not* be easy to model the ways that the archaeological record came about or to predict where archaeological materials in general, or specific sorts of significant archaeological materials, will be found. Claims that predictive modeling is easy or that a particular model is highly successful should be carefully examined in light of this chapter. Does the model in question consider past systems organization? Are empirical "predictive models" of general utility not only in predicting the locations of archaeological materials but in explaining the systemic mechanisms behind them? If not, they are likely not to be generally successful and applicable, for mechanisms must be elucidated before their consequences can be determined.

We are presently at a very crucial point in archaeological science and in the practice of cultural resource management. Management requires that we be able to predict the locations and significance of archaeological resources, and archaeology must discover how to do this. Fulfilling this goal will require concentrated and dedicated research that may not, at all times, appear to be totally directed toward the pursuit of simply identifying and conserving sites. Management must be

patient and supportive of the genuine pursuit of archaeological explanation, for it is only through explanation that we can understand *anything* about the past through the archaeological record. Archaeological prediction is a new frontier, and all aspects of it must be justified and proven in explanatory terms.

We would first like to thank LuAnn Wandsnider, Eileen Camilli, Bryan Marozas, Signa Larralde, Lewis Binford, Michael Schiffer, Rob Foley, Jim Hester, Frank McManamon, Robert Dunnell, and Max Ayer. Discussions and correspondence with these and other professionals over the past few years have helped shape our thinking on the subject of predictive modeling and the organization of past human behavior. Dan Martin, Jim Judge, Lynne Sebastian, and Chris Kincaid were instrumental in bringing this volume to fruition, and we thank them as well. Most of all, the senior author adds that Lynne and June-el have dealt with horribly complex sentences (they increase in complexity as the thoughts they express do), non sequiturs, and worst of all, bizarre references. To the technical editing team more than to any others we owe the sequentiality and readability of this chapter.

The senior author participated in a symposium at the 1984 meeting of the Society for American Archaeology in Portland in which he delivered, with LuAnn Wandsnider and Signa Larralde, a paper entitled "Predictive Modeling: Current Abuses of the Archaeological Record and Prospects for Explanation." This symposium was chaired by Barry Holt, who should also be thanked. It was shortly after that meeting that the contract resulting in our participation in this volume was awarded. The papers in that symposium were perceptive, for their day, but they portrayed archaeological predictive modeling as a panacea for the manager—a means by which survey and excavation could be avoided.

This same author was gratified to see that at the 1988 meeting of the Society for American Archaeology in Phoenix, papers focusing on predictive modeling were no longer heralding this technique as a cure-all for archaeological survey obligations. Both government-employed managers and university researchers have now accepted "predictive modeling" for what it should be: a research tool through which we can test our ideas about the past against the data provided by the archaeological record. In a very real sense, predictive modeling doesn't predict the archaeological record—the archaeological record predicts what we learn through predictive modeling. Predictive modeling allows us to refine our ideas, and the computer methods we have for expressing them, in light of what archaeologists have found and continue to find on and in the earth.

Nothing better points out the value and applications of predictive modeling than this logical progression within the last few years. As pointed out in our chapter, modeling is an interactive procedure by which archaeologists and managers learn about what it is they want to know about the past, and how this is expressed and verified (or perhaps not) by the archaeological record. It is in a very real sense an embodiment of the process of archaeological science. In archaeology, the process of science cannot be separated from the process of management. The two are inherently related. In light of this assertion, we would also like to acknowledge the efforts of all those "archaeological managers" who have shown interest in archaeological predictive modeling, whatever their rationales and goals. They are the people to whom this chapter is dedicated, and to whom it is directed. We express the hope that we can work alongside them in years to come in refining the complex and exciting methods we'll all want and need to perfect in order to arrive at meaningful depictions of past human behavior and the value of our knowledge of these in the present.

The junior author would also like to acknowledge the support of the School of American Research in Santa Fe, where he was residing and working when he made the changes in the text in response to the review comments.

REFERENCES CITED

Allchin, Bridget
1966 *The Stone-Tipped Arrow: The Late Stone-Age Hunters of the Tropical Old World.* Barnes and Noble, New York.

Artz, Joe Alan, and Kenneth C. Reid
1983 Modeling the Location of Buried Archaic Sites in Northeastern Oklahoma: A Geomorphic Perspective. Paper presented at the 48th Annual Meeting of the Society for American Archaeology, Pittsburgh.

Baker, Craig, and Steven E. Sessions
1979 The Environmental Context of Archaeological Sites in the EPCC Survey Area. In *The Archaeology of Southwest Gallegos Mesa: The EPCC Survey Project,* edited by Steven E. Sessions. Navajo Nation Papers in Anthropology 1. Window Rock, Arizona.

Behrensmeyer, Anna K., and Andrew P. Hill, editors
1980 *Fossils in the Making.* University of Chicago Press, Chicago.

Bettinger, Robert L.
1977a Predicting the Archaeological Potential of the Inyo-Mono Region of Eastern California. In *Conservation Archaeology,* edited by Michael B. Schiffer and George J. Gumerman, pp. 217–226. Academic Press, New York.

1977b *The Surface Archaeology of the Long Valley Caldera, Mono County, California.* Archaeological Research Unit Monographs No. 1. University of California, Riverside.

Bettinger, Robert L., and Martin A. Baumhoff
1982 The Numic Spread: Great Basin Cultures in Competition. *American Antiquity* 47:485–503.

Binford, Lewis R.
1965 Archaeological Systematics and the Study of Culture Process. *American Antiquity* 31:203–210.

1976 Forty-seven Trips: A Case Study in the Character of Some Formation Processes of the Archaeological Record. In *Contributions to Anthropology: The Interior Peoples of Northern Alaska,* edited by E. S. Hall, Jr., pp. 299–351. National Museum of Man, Mercury Series No. 49. Ottawa.

1978 *Nunamiut Ethnoarchaeology.* Academic Press, New York.

1979 Organization and Formation Processes: Looking at Curated Technologies. *Journal of Anthropological Research* 35:255–273.

1980 Willow Smoke and Dogs' Tails: Hunter-Gatherer Settlement Systems and Archaeological Site Formation. *American Antiquity* 45:1–17.

1981 Behavioral Archaeology and the "Pompeii Premise." *Journal of Anthropological Research* 37:195–208.

1982 The Archaeology of Place. *Journal of Anthropological Archaeology* 1:5–13.

1983 *In Pursuit of the Past: Decoding the Archaeological Record.* Thames and Hudson, London.

Binford, Lewis R., editor
1977 *For Theory Building in Archaeology.* Academic Press, New York.

Binford, Lewis R., and Jeremy A. Sabloff
1982 Paradigms, Systematics, and Archaeology. *Journal of Anthropological Research* 38:137–153.

Boserup, Esther
1965 *The Conditions of Agricultural Growth: The Economics of Agrarian Change Under Population Pressure.* Aldine, Chicago.

Bowers, Peter M., Robson Bonnichsen, and David M. Hock
1983 Flake Dispersal Experiments: Non-Cultural Transformation of the Archaeological Record. *American Antiquity* 48:553–572.

Brain, C. K.
1967a Bone Weathering and the Problem of Bone Pseudo-Tools. *South African Journal of Science* 63:97–99.

1967b Hottentot Food Remains and Their Bearing on the Interpretation of Fossil Bone Assemblages. *Scientific Papers of the Namib Desert Research Station* 32:1–11.

1969 The Contribution of Namib Desert Hottentots to an Understanding of *Australopithecine* Bone Accumulations. *Scientific Papers of the Namib Desert Research Station* 39:13–22.

1981 *The Hunters or the Hunted? An Introduction to African Cave Taphonomy.* University of Chicago Press, Chicago.

Buechner, Helmuth K.
1971 The Eco-System Level of Organization. In *Hierarchically Organized Systems in Theory and Practice,* edited by Paul A. Weiss, pp. 45–58. Hafner, New York.

Bunn, Henry, John W. K. Harris, Glynn Isaac, Zefe Kaufulu, Ellen Kroll, Kathy Schick, Nicholas Toth, and Anna K. Behrensmeyer
1980 FxJj50: An Early Pleistocene Site in Northern Kenya. *World Archaeology* 12(2):109–136.

Butzer, K. W.
1977 Geo-Archaeology in Practice. *Reviews in Archaeology* 4:125–131.

1982 *Archaeology as Human Ecology: Method and Theory for a Contextual Approach.* Cambridge University Press, Cambridge, England.

Camilli, Eileen L., LuAnn Wandsnider, and James I. Ebert
1988 *Distributional Survey and Excavation of Archaeological Landscapes in the Vicinity of El Paso, Texas.* Bureau of Land Management. Government Printing Office, Washington, D.C., in press.

Carr, Christopher
1981 The Polythetic Organization of Archaeological Tool Kits and an Algorithm for Defining Them. Paper presented at the 46th Annual Meeting of the Society for American Archaeology, San Diego.

1984 The Nature of Organization of Intrasite Archaeological Records and Spatial Analytic Approaches to Their Investigation. In *Advances in Archaeological Method and Theory,* vol. 7, edited by Michael B. Schiffer, pp. 103–222. Academic Press, New York.

Chandler, Susan M., and Paul R. Nickens
1983 Regional Sampling and Predictive Modeling in CRM: Some Examples from the Northern Colorado Plateau. Paper presented at the 48th Annual Meeting of the Society for American Archaeology, Pittsburgh.

Clarke, D. L.
1968 *Analytical Archaeology.* Methuen, London.

Colwell, Robert K.
1974 Predictability, Constancy, and Contingency of Periodic Phenomena. *Ecology* 55:1148–1153.

Davis, R. P. Stephen, Jr.
1980a A Summary Report of Probabilistic "Non-Site" Sampling in Tellico Reservoir, 1979. In *The 1979 Archaeological and Geological Investigations in the Tellico Reservoir,* edited by Jefferson Chapman, pp. 59–60. Report of Investigations No. 29. Department of Anthropology, University of Tennessee, Nashville.

1980b Archaeological Survey in the Lower Little Tennessee River Valley: A Comparison of Probabilistic Sampling Strategies. Paper presented at the 37th Annual Meeting of the Southeast Archaeological Conference, New Orleans.

Drager, Dwight L., and Arthur K. Ireland, editors
1986 *The Seedskadee Project: Remote Sensing in Non-Site Archaeology.* Branch of Remote Sensing, Southwest Cultural Resources Center, National Park Service, Albuquerque, and Bureau of Reclamation, Salt Lake City.

Dunnell, Robert C., and William S. Dancey
1983 The Siteless Survey: A Regional Scale Data Collection Survey. In *Advances in Archaeological Method and Theory,* vol. 6, edited by Michael B. Schiffer, pp. 267–287. Academic Press, New York.

Ebert, James I.
1983a Request for Proposal No. 7029-3-0015: Cultural Resources Sample Survey in the Seedskadee Cultural Resources Assessment Project Area, Southwestern Wyoming. National Park Service, Santa Fe. On file, Branch of Remote Sensing, Southwest Cultural Resources Center, National Park Service, University of New Mexico, Albuquerque.

1983b Seedskadee Project: Proposal for Remote Sensing and Class I Survey. Ms. on file, Branch of Remote Sensing, Southwest Cultural Resources Center, National Park Service, University of New Mexico, Albuquerque.

1984 Remote Sensing Applications in Archaeology. In *Advances in Archaeological Method and Theory,* vol. 7, edited by Michael B. Schiffer, pp. 293–362. Academic Press, New York.

1986 *Distributional Archaeology: Nonsite Discovery, Recording and Analytical Methods for Application to the Surface Archaeological Record.* Ph.D. dissertation, Department of Anthropology, University of New Mexico, Albuquerque.

Ebert, James I., Signa Larralde, and LuAnn Wandsnider
1983 Distributional Archaeology: Survey, Mapping, and Analysis of Surface Archaeological Materials in the Green River Basin, Wyoming. Paper presented at the 41st Plains Conference, Rapid City, South Dakota.

1984 Predictive Modeling: Current Abuses of the Archaeological Record and Prospects for Explanation. Paper presented at the 49th Annual Meeting of the Society for American Archaeology, Portland.

Ellen, Roy
1982 *Environment, Subsistence, and System: The Ecology of Small-Scale Social Formations.* Cambridge University Press, Cambridge, England.

Ellis, Florence H.
1978 Small Structures Used by Historic Pueblo Peoples and Their Immediate Ancestors. In *Limited Activity and Occupation Sites: A Collection of Conference Papers,* edited by Albert E. Ward, pp. 59–68. Contributions to Anthropology No. 1. Center for Anthropological Studies, Albuquerque.

Fish, Suzanne K., and Paul R. Fish
1978 The Small House in a Social Network: Some Mesoamerican Examples. In *Limited Activity and Occupation Sites: A Collection of Conference Papers,* edited by Albert E. Ward, pp. 51–59. Contributions to Anthropology No. 1. Center for Anthropological Studies, Albuquerque.

Foley, Robert A.
1977 Space and Energy: A Method for Analyzing Habitat Value and Utilization in Relation to Archaeological Sites. In *Spatial Archaeology,* edited by David L. Clarke, pp. 163–187. Academic Press, New York.

1978 Incorporating Sampling into Initial Research Design: Some Aspects of Spatial Archaeology. In *Sampling in Contemporary British Archaeology,* edited by J. Cherry, C. Gamble, and S. Shennan, pp. 49–65. British Archaeological Reports, British Series 50. Oxford.

1980 The Spatial Component of Archaeological Data: Off-Site Methods and Some Preliminary Results from the Amboseli Basin of Southern Kenya. *Proceedings of the VIII Pan-African Congress in Prehistory and Quaternary Studies,* pp. 39–40.

1981a Off-Site Archaeology: An Alternative Approach for the Short-Sited. In *Pattern of the Past: Essays in Honor of David L. Clarke*, edited by I. Hodder, G. Isaac, and N. Hammond, pp. 157–183. Cambridge University Press, Cambridge, England.

1981b A Model of Regional Archaeological Structure. *Proceedings of the Prehistoric Society* 47:1–17.

1981c *Off-Site Archaeology and Human Adaptation in Eastern Africa: An Analysis of Regional Artefact Density in the Amboseli, Southern Kenya.* B.A.R. International Series 97, Cambridge Monographs in African Archaeology 3. British Archaeological Reports 265. Oxford.

Fowler, Don D., and Catherine S. Fowler, editors

1971 *Anthropology of the Numa: John Wesley Powell's Manuscripts on the Numic Peoples of Western North America, 1868–1880.* Smithsonian Contributions to Anthropology No. 14. Washington, D. C.

Gifford, Diane P.

1977a *Observations of Modern Human Settlements as an Aid to Archaeological Interpretation.* Ph.D. dissertation, Department of Anthropology, University of California, Berkeley.

1977b Ethnoarchaeological Observations on Natural Processes Affecting Cultural Materials. In *Explorations in Ethnoarchaeology*, edited by R. A. Gould, pp. 77–102. University of New Mexico Press, Albuquerque.

1980 Ethnoarchaeological Contributions to the Taphonomy of Human Sites. In *Fossils in the Making*, edited by Anna K. Behrensmeyer and Andrew P. Hill, pp. 94–107. University of Chicago Press, Chicago.

1981 Taphonomy and Paleoecology: A Critical Review of Archaeology's Sister Discipline. In *Advances in Archaeological Method and Theory*, vol. 4, edited by Michael B. Schiffer, pp. 365–438. Academic Press, New York.

Gifford, Diane P., and Anna K. Behrensmeyer

1977 Observed Formation and Burial of a Recent Human Occupation Site in Kenya. *Quaternary Research* 8:245–266.

Gladfelter, B. G.

1977 Geoarchaeology: The Geomorphologist and Archaeology. *American Antiquity* 42:519–538.

Gould, Richard A.

1980 *Living Archaeology.* Cambridge University Press, Cambridge, England.

Harako, Reizo

1978 Patterns of Mobility Among Hunting and Gathering People—An Ecological Comparative Study of Bushmen and Pygmies. *Reports for "Mobility for Man and Society," IATSS Symposium on Traffic Science*, pp. 111–118. Tokyo.

Harpending, Henry, and Herbert Davis

1977 Some Implications for Hunter-Gatherer Ecology Derived from the Spatial Structure of Resources. *World Archaeology* 8(3):277–286.

Harris, Marvin D.

1977 *Cannibals and Kings: The Origins of Culture.* Random House, New York.

Hitchcock, Robert K., and James I. Ebert

1984 Foraging and Food Production among Kalahari Hunter-Gatherers. In *From Hunters to Farmers: The Causes and Consequences of Food Production in Africa*, edited by J. Desmond Clark and Steven A. Brandt, pp. 328–348. University of California Press, Berkeley.

House, J. H., and D. L. Ballenger

1976 *An Archaeological Survey of Interstate 77 Route in the South Carolina Piedmont.* Research Manuscript Series No. 104. Division of Advanced Studies and Research, Institute of Archaeology and Anthropology, University of South Carolina, Columbia.

Isaac, Glynn Ll.

1966 New Evidence from Olorgesailie Relating to the Character of Acheulian Occupation Sites. In *Actas del V Congreso Panafricano de Prehistoria y de Estudio del Cuaternario II*, edited by L. D. Cuscoy, pp. 135-145. Publicaciones del Museo Arqueologico Santa Cruz de Tenerife 6.

1967 Towards the Interpretation of Occupation Debris: Some Experiments and Observations. *Kroeber Anthropological Society Papers* 37:31-57.

1978 Casting the Net Wide: A Review of Archaeological Evidence for Early Hominid Land-Use and Ecological Relations. In *Current Arguments on Early Man*, edited by Lars-Konig Konigsson, pp. 226-251. Pergamon Press, Oxford.

1981 Stone Age Visiting Cards: Approaches to the Study of Early Land Use Patterns. In *Pattern of the Past: Essays in Honor of David Clarke*, edited by I. Hodder, G. Isaac, and N. Hammond, pp. 131-156. Cambridge University Press, Cambridge, England.

Isaac, Glynn Ll., and J. W. K. Harris

1975 The Scatter Between the Patches. Paper presented to the Kroeber Anthropological Society, May, Berkeley.

1978 Archaeology. In *The Fossil Hominids and an Introduction to Their Context, 1968-1974*, edited by M. G. Leakey and R. E. Leakey, pp. 64-85. Koobi Fora Research Project, vol I. Clarendon Press, Oxford.

Isaac, Glynn Ll., J. W. K. Harris, and Fiona Marshall

1981 Small is Informative: The Application of the Study of Mini-Sites and Least Effort Criteria in the Interpretation of the Early Pleistocene Archaeological Record at Koobi Fora, Kenya. *Union Internacional de Ciencias Prehistoricas y Protohistoricas, X Congreso*, pp. 102-119. UNESCO, Mexico.

Jochim, Michael A.

1976 *Hunter-Gatherer Subsistence and Settlement: A Predictive Model.* Academic Press, New York.

Judge, W. James

1973 *Paleoindian Occupation of the Central Rio Grande Valley in New Mexico.* University of New Mexico Press, Albuquerque.

1981 Transect Sampling in Chaco Canyon—Evaluation of a Survey Technique. In *Archaeological Surveys of Chaco Canyon*, edited by Alden C. Hayes, David M. Brugge, and W. James Judge, pp. 107-137. National Park Service, Washington, D.C.

Kelley, Klara B., Stephen C. Lent, John C. Acklen, and Jeffrey L. Boyer

1982 Analysis of Data from the Navajo Site Reports. In *Anasazi and Navajo Land Use in the McKinley Mine Area near Gallup, New Mexico*, vol. I (Part One), edited by Christina G. Allen and Ben A. Nelson, pp. 413-446. Office of Contract Archeology, University of New Mexico, Albuquerque.

Kelly, Robert L.

1983 Hunter-Gatherer Mobility Strategies. *Journal of Anthropological Research* 29:277-306.

Kemrer, Meade F., editor

1982 *Archaeological Variability Within the Bisti-Star Lake Region, Northwestern New Mexico.* ESCA-Tech Corporation. Submitted to the Bureau of Land Management, Contract No. YA553-CTO-114. Copies available from the Bureau of Land Management, Albuquerque District Office.

Kintigh, K. W., and A. J. Ammerman

1982 Heuristic Approaches to Spatial Analysis in Archaeology. *American Antiquity* 47:31-63.

Kirkby, A., and M. J. Kirkby

1976 Geomorphological Processes and the Surface Survey of Archaeological Sites in Semi-Arid Areas. In *Geoarchaeology: Earth Sciences and the Past*, edited by D. A. Davidson and M. L. Shackley, pp. 229-253. Duckworth, London.

Klesert, Anthony L.
1983 A Predictive Model of Site Potential in the Rio Grande National Forest, Colorado. Paper presented at the 48th Annual Meeting of the Society for American Archaeology, Pittsburgh.

Kohler, Timothy A.
1976 The Fourth Law of Thermodynamics in Cultural Evolution. *Florida Journal of Anthropology* 1(2):26–32.

Kohler, Timothy A., and Eric Blinman
1987 Solving Mixture Problems in Archaeology: Analysis of Ceramic Materials for Dating and Demographic Reconstruction. *Journal of Anthropological Archaeology* 6:1–28.

Kohler, Timothy A., and Meredith Matthews
1988 Long-Term Anasazi Land Use and Forest Reduction: A Case Study from Southwest Colorado. *American Antiquity*, in press.

Kohler, Timothy A., and Sandra C. Parker
1986 Predictive Models for Archaeological Resource Location. In *Advances in Archaeological Method and Theory*, vol. 9, edited by Michael B. Schiffer, pp. 397–452. Academic Press, New York.

Kohler, Timothy A., Carl J. Phagan, and E. Blinman
1985 *Sources of Confounding Variability in Archaeological Collections.* Dolores Archaeological Project Report No. 282. Dolores, Colorado.

Kvamme, Kenneth
1980 Predictive Model of Site Location in the Glenwood Springs Resource Area. In *Class II Cultural Resource Inventory of the Glenwood Springs Resource Area, Grand Junction District, Colorado.* Submitted to the Bureau of Land Management, Grand Junction District Office, Colorado.

1982 Methods for Analyzing and Understanding Hunter-Gatherer Site Location as a Function of Environmental Variation. Paper presented at the 47th Annual Meeting of the Society for American Archaeology, Minneapolis.

1983a A Computer Approach for Wide-Area Probability Surface Mapping of Archaeological Site Presence. Paper presented at the 48th Annual Meeting of the Society for American Archaeology, Pittsburgh.

1983b A Manual for Predictive Site Location Models: Examples from the Grand Junction District, Colorado. Draft submitted to the Bureau of Land Management, Grand Junction District, Colorado.

Larralde, Signa
1984 Quality Control in Lithics Analysis: A Test of Precision. *Haliksa'i, University of New Mexico Contributions to Anthropology* 3(Spring):1–8. Albuquerque.

Lewarch, Dennis E., and Michael J. O'Brien
1981 The Expanding Role of Surface Assemblages in Archaeological Research. In *Advances in Archaeological Method and Theory*, vol. 4, edited by Michael B. Schiffer, pp. 297–342. Academic Press, New York.

Limp, W. Frederick
1983 Location Choice and Settlement Predictions: The Conformity of Method and Theory. Paper presented at the 48th Annual Meeting of the Society for American Archaeology, Pittsburgh.

Lynch, B.
1980 Site Artifact Density and the Effectiveness of Shovel Probes. *Current Anthropology* 21:516–517.

Marrinan, Rochelle A.
1975 *Ceramics, Molluscs, and Sedentism: The Late Archaic Period on the Georgia Coast.* Ph.D. dissertation, University of Florida. University Microfilms, Ann Arbor.

Mazel, Aron, and John Parkington
1981 Stone Tools and Resources: A Case Study from Southern Africa. Ms. on file, Department of Anthropology, University of Cape Town, South Africa.

McAllister, Shirley Powell, and Fred Plog
1978 Small Sites in the Chevelon Drainage. In *Limited Activity and Occupation Sites: A Collection of Conference Papers,* edited by Albert E. Ward, pp. 17–24. Contributions to Anthropology No. 1. Center for Anthropological Studies, Albuquerque.

McManamon, Francis P.
1981a Parameter Estimation and Site Discovery in the Northeast. *Contract Abstracts and CRM Archaeology* 1(3):43–48.

1981b Probability Sampling and Archaeological Survey in the Northeast: An Estimation Approach. In *Foundations of Northeast Archaeology,* edited by D. R. Snow, pp. 195–227. Academic Press, New York.

Moore, Bruce M.
1978 Are Pueblo Field Houses a Function of Urbanization? In *Limited Activity and Occupation Sites: A Collection of Conference Papers,* edited by Albert E. Ward, pp. 9–16. Contributions to Anthropology No. 1. Center for Anthropological Studies, Albuquerque.

Nance, C. Roger, Harry Holstein, and David C. Hurst
1983 Evaluation of Multiple Regression Models Predicting Archaeological Site Distributions at Fort McClellan, Alabama. Paper presented at the 48th Annual Meeting of the Society for American Archaeology, Pittsburgh.

Nance, J. D.
1980 Non-Site Sampling in the Lower Cumberland River Valley, Kentucky. *Midcontinental Journal of Archaeology* 5:169–192.

1981 Statistical Fact and Archaeological Faith: Two Models in Small Site Sampling. *Journal of Field Archaeology* 8:151–165.

Oswalt, Wendell H.
1973 *Habitat and Technology: The Evolution of Hunting.* Holt, Rinehart and Winston, New York.

Parker, Sandra
1983 Predictive Modeling of Site Settlement Systems Using Multivariate Logistics. In *For Concordance in Archaeological Analysis: Bridging Data Structure, Quantitative Technique, and Theory,* edited by Christopher Carr, pp. 173–207. Westport, Kansas City.

Peebles, Thomas C.
1983 Discriminant Analysis in Site Predictive Models: Problems and Prospects. Paper presented at the 48th Annual Meeting of the Society for American Archaeology, Pittsburgh.

Piaget, Jean
1978 *Behavior and Evolution,* translated by Donald Nicholson-Smith. Pantheon Books, New York.

Powell, Shirley Louise
1980 *Material Culture and Behavior: A Prehistoric Example from the American Southwest.* Ph.D. dissertation, Department of Anthropology, Arizona State University, Tempe.

Reynolds, William E., editor
1983 *The Archaeology and History of the La Barge Area, Southwestern Wyoming: A Class II Cultural Resources Inventory.* ESCA-Tech Corporation. Submitted to the Bureau of Land Management, Rock Springs District Office, Wyoming.

Russell, Scott C.
1978 The Agricultural Field House: A Navajo Limited Occupation and Special Use Site. In *Limited Activity and Occupation Sites: A Collection of Conference Papers,* edited by Albert E. Ward, pp. 35–40. Contributions to Anthropology No. 1. Center for Anthropological Studies, Albuquerque.

Sabo, George, III, and David Waddell
1983 Adaptation Type Models and Site Prediction in the Arkansas Ozarks. Paper presented at the 48th Annual Meeting of the Society for American Archaeology, Pittsburgh.

Sahlins, Marshall D.
1972 *Stone Age Economics.* Travistock, London.

Schiffer, Michael B.
1972 Archaeological Context and Systemic Context. *American Antiquity* 37:156–165.

1983 Toward the Identification of Formation Processes. *American Antiquity* 48:675–706.

Scholtz, S. C.
1980 Predictive Models and Survey Strategy in the Sparta Mine Area. Paper presented at the 37th Annual Southeastern Archaeological Conference, New Orleans.

1981 Location Choice in Sparta. In *Settlement Prediction in Sparta,* by R. F. Lafferty, J. L. Otinger, S. C. Scholta, W. F. Limp, B. Watkins, and R. D. Jones, pp. 207–222. Arkansas Archaeological Survey Research Series No. 14. Fayetteville.

Simek, Jan F., and Roy R. Larick
1983 The Recognition of Multiple Spatial Patterns: A Case Study from the French Upper Paleolithic. *Journal of Archaeological Science* 10:165–180.

Smith, Anne B.
1974 *Ethnography of the Northern Utes.* Papers in Anthropology 17. Museum of New Mexico, Santa Fe.

Spurling, B.
1980 Site Discovery and Assessment Techniques for Mixed Cover Survey Regimes. *Saskatchewan Archaeology* 1:25–56.

Stahle, David W., and James E. Dunn
1982 An Analysis and Application of the Size Distribution of Waste Flakes from the Manufacture of Bifacial Stone Tools. *World Archaeology* 14:84–97.

Steward, Julian H.
1938 *Basin-Plateau Aboriginal Sociopolitical Groups.* Bureau of American Ethnology Bulletin 120. Smithsonian Institution, Washington, D.C.

Tanaka, Jiro
1969 The Ecology and Social Structure of Central Kalahari Bushmen: A Preliminary Report. *Kyoto University African Studies* 3:1–26.

1976 Subsistence Ecology of the Central Kalahari San. In *Kalahari Hunter-Gatherers: Studies of the !Kung San and Their Neighbors,* edited by Richard B. Lee and Irven DeVore, pp. 98–119. Harvard University Press, Cambridge.

Thomas, David Hurst
1973 An Empirical Test of Steward's Model of Great Basin Settlement Patterns. *American Antiquity* 38:155–176.

1975 Nonsite Sampling in Archaeology: Up the Creek Without a Site? In *Sampling in Archaeology,* edited by James W. Mueller, pp. 61–81. University of Arizona Press, Tucson.

1983 *The Archaeology of Monitor Valley, vol. 1: Epistemology.* Anthropological Papers of the American Museum of Natural History 58(1). New York.

Torrence, Robin
1983 Time Budgeting and Hunter-Gatherer Technology. In *Hunter-Gatherer Economy in Prehistory,* edited by G. Bailey, pp. 11–22. Cambridge University Press, Cambridge.

Vita-Finzi, E., and E. S. Higgs
1970 Prehistoric Economy in the Mount Carmel Area of Palestine: Site Catchment Analysis. *Proceedings of the Prehistoric Society* 36:1–37.

Waddell, David B.
1983 Selection of Attributes for Predictive Modeling of Site Location. Paper presented at the 48th Annual Meeting of the Society for American Archaeology, Pittsburgh.

Wandsnider, LuAnn, and James I. Ebert
1983 Modeling Climatic and Landform Factors Affecting the Character of the Archaeological Record in Arid Lands: A Remote Sensing Approach in Southwestern Archaeology. Paper presented at the Silver Anniversary Meeting of the Western Social Science Association for Arid Lands Studies, Albuquerque.

1984 Accuracy of Archaeological Survey in the Seedskadee Project Area, Southwestern Wyoming. *Haliksa'i, University of New Mexico Contributions to Anthropology* 3(Spring):9–21. Albuquerque.

Wandsnider, LuAnn, and Signa Larralde
1984 Seedskadee Cultural Resources Assessment Project: Report to the Branch of Remote Sensing, National Park Service. Chambers Consultants and Planners, Albuquerque. Submitted to the Branch of Remote Sensing, Southwest Cultural Resource Center, National Park Service, Albuquerque.

Warren, R. E.
1979 Archaeological Site Survey. In *Cannon Reservoir Human Ecology Project—A Regional Approach to Cultural Continuity and Change,* edited by M. J. O'Brien and R. Warren, pp. 71–100. Interim report submitted to the U.S. Army Corps of Engineers, St. Louis District, Missouri.

Weiss, Paul A.
1973 *The Science of Life: The Living System—A System for Living.* Futura, Mt. Kisco, New York.

Whallon, Robert
1973 Spatial Analysis of Occupation Floors, I: Application of Dimensional Analysis of Variance. *American Antiquity* 38:320–328.

1974 Spatial Analysis of Occupation Floors, II: Application of Nearest Neighbor Analysis. *American Antiquity* 39:16–34.

1984 Unconstrained Clustering for the Analysis of Spatial Distributions in Archaeology. In *Intrasite Spatial Analysis in Archaeology,* edited by H. J. Hietala and P. Larson, pp. 242–277. Cambridge University Press, Cambridge, England.

Wilkinson, R. G.
1973 *Poverty and Progress: An Ecological Model of Economic Development.* Methuen, London.

Willey, Gordon R., and Jeremy A. Sabloff
1974 *A History of Archaeology.* W. H. Freeman, San Francisco.

Wills, W. H.
1980 Ethnographic Observation and Archaeological Interpretation: The Wikmunkan of Cape York Peninsula, Australia. In *The Archaeological Correlates of Hunter-Gatherer Societies: Studies from the Ethnographic Record,* edited by F. E. Smiley, Carla M. Sinopoli, H. Edwin Jackson, W. H. Wills, and S. A. Gregg, pp. 76–99. Michigan Discussions in Anthropology 5(1–2). University of Michigan, Ann Arbor.

Winterhalder, Bruce
1980 Environmental Analysis in Human Evolution and Adaptation Research. *Human Ecology* 8:135–170.

Wood, W. R., and D. L. Johnson
1978 A Survey of Disturbance Processes in Archaeological Site Formation. In *Advances in Archaeological Method and Theory,* vol. 1, edited by Michael B. Schiffer, pp. 315–381. Academic Press, New York.

Yellen, John E.
1976 Settlement Patterns of the !Kung: An Archaeological Perspective. In *Kalahari Hunter-Gatherers: Studies of the !Kung San and Their Neighbors*, edited by Richard B. Lee and Irven DeVore, pp. 47–72. Harvard University Press, Cambridge.

Chapter 5

AN OVERVIEW OF STATISTICAL METHOD AND THEORY FOR QUANTITATIVE MODEL BUILDING

Martin R. Rose and Jeffrey H. Altschul

This chapter focuses on some of the conceptual aspects of building quantitative predictive models. The discussion is aimed primarily at correlative models, although many of the topics addressed will also apply to other types of models. It must be emphasized that the aim is not a "cookbook" of statistical procedures involved in producing a predictive model. There are many textbooks devoted to univariate, bivariate, and multivariate statistics; some even emphasize specific predictive modeling techniques, such as multiple and logistic regression and discriminant function analysis. For the most part, the reader will be referred to these textbooks for detailed discussions of the nuts and bolts of specific procedures, although it is impossible to avoid including some formulas and detailed discussions in this chapter. Statistical techniques, especially multivariate ones, are not simple procedures. Some may be relatively easy to understand from a conceptual standpoint, but many of the assumptions and intricacies of the procedures are not amenable to a cursory examination. If statistical techniques are going to be used to build a predictive model, the researcher must be willing to invest the time to learn how to do it correctly.

This chapter begins by defining the problem that predictive modeling attempts to address—namely, the distribution of sites in space. Specifically, predictions about site locations in a region are generated on the basis of observed associations between a set of independent predictor variables and site locations in a sample of locations in that region. This information about the attributes of site location in the sample is used to "predict" site location attributes for an area not included in the original sample but for which observations for the same set of independent variables have been made.

Once the nature of the problem has been defined, we will consider the nature of variables, scales, and distributions appropriate to different discrete and continuous random variables. Variables should be designed to measure certain theoretical aspects of the phenomena of interest, and further, each variable ideally should reflect only one dimension of variability. The different scales of measurement commonly employed in statistics will be briefly reviewed so that the limitations

they place on potential analytical procedures can be noted. Finally, different types of discrete and continuous probability distributions will be examined, and the implications of these distributions for parametric vs nonparametric statistics will be outlined.

Statistical description and inference in the model-building process are examined in the third major section of the chapter. Two distinct statistical parts to the model-building process are recognized. The first is the definition and measurement of the associations among one or more independent variables and of their relationships to site location. The second concerns generalizing from these sample-based associations to the larger population. Some general univariate descriptive statistics and bivariate and multivariate tests of association will be described. We will highlight some of the concepts involved in the use of inferential statistics and emphasize the probability-based nature of these statistics and their dependence on some type of probabilistic sampling scheme.

A large part of the chapter is devoted to the topics of defining differences among groups and producing some type of classification of these groups. The first objective of the analytical procedure is to take two or more known groups defined on the basis of a set of independent variables and then determine which of the variables provides the most important discrimination between or among groups. The second goal is to capitalize mathematically on the group differences and produce a function or set of functions that allow the classification of unknown cases into the most likely group. The importance of "cleaning up the data" is discussed as a necessary early step in predictive modeling. This process helps to provide a set of variables that can be used to distinguish site distributions from nonsites or different types of sites from one another. The requirement for homogeneous groups is described, and factors that work against homogeneity, such as temporal and functional variability in sites, the difficulty of defining site classes using cultural resource management data, and the inherent heterogeneity of nonsite points, are discussed. Attention is given to appropriate methods for defining site classes and reducing heterogeneity in data sets using the techniques of cluster analysis and principal components/factor analyses. Finally, three of the more popular techniques for assigning cases to groups in predictive models—general linear regression, logistic regression, and discriminant function analysis—are described.

The subject of the final section of the chapter is the validation and generalization of predictive models. Three different perspectives are suggested for model validation: the use of independent data, split-sample validation procedures, and the use of synthetic (simulated) multivariate data sets. Model generalization using computer-generated contour maps is discussed briefly.

MODELING SITE LOCATION

The Problem

Predictive modeling is based on the assumptions that human behavior is patterned and that the outcomes of the decisions that people make about where and where not to live are also patterned. Most forms of spatial analysis also assume that the patterns in settlement behavior can be discerned by studying the locations of sites. Scientists studying modern patterns of settlement can examine all the components of the system, but archaeologists are restricted to the patterns that can be discerned from partial, and often biased, data. Predictive modeling further assumes that patterns of locational behavior in a particular region can be perceived through statistical analyses of samples drawn from the archaeological record, and that the resulting patterns can be generalized to the larger area.

From a quantitative standpoint, predictive modeling is a process that permits us to determine the long-term relative frequency, or probability, that any particular location within a region contains a site. For the purposes of this discussion we will sidestep the problem of varying site sizes and will assume that a *location,* defined here as some areal unit (e.g., hectare, acre, square kilometer, etc.) contains either one site or no sites.

A diagram can be used to illustrate this situation. In Figure 5.1 the space within the borders of the rectangle represents the region in question—the area covered by a national forest, for example, or by a coal-lease tract. The dots in the diagram represent site locations. If our purpose is to determine the probability that a location of specified size will contain a site, one approach would be to divide the region into units of the specified size, survey each unit, and then tabulate the results. For example, we might divide the space in Figure 5.1 into 500 units, each representing some specified area. If we were to inventory the entire space and find that 10 units contained sites, we could calculate a proportion of 10/500 or 0.02 sites per unit. With no other information available we might take this proportion to be the probability of finding a site in a unit selected by chance.

This example highlights two important points about predictive modeling. First, how is the probability of an event determined? If it were necessary to survey an entire region in order to determine the probability of finding a site in a location, there would be no need for predictive modeling. Fortunately this is not the case. In explaining why this is so, we must introduce the concept of a random experiment. Put simply, a random experiment involves certain actions conducted under specified conditions that has as its outcome one (and only one) of a set of possible results (usually termed *simple results*). Before the experiment is conducted we have no way of predicting which simple result will occur.

Returning to our hypothetical archaeological survey, we can construct a random experiment in which every time a unit is surveyed there are two possible simple results: "yes," a site is present, and "no," a site is not present. If we are

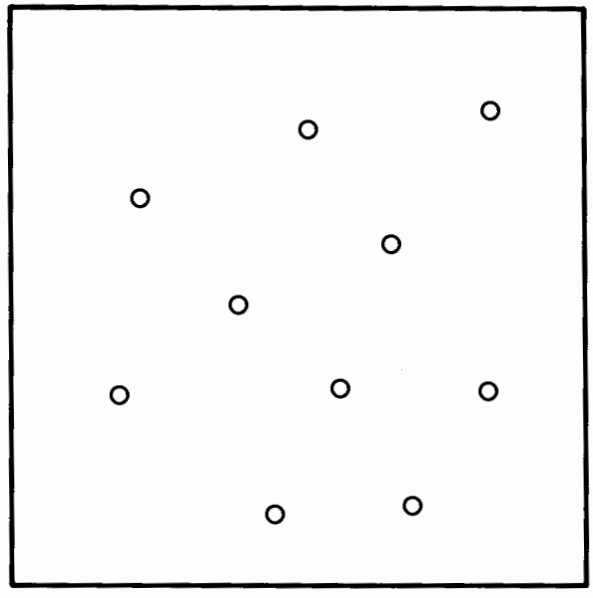


Figure 5.1. Diagram of a region (dots represent site locations).

interested in those cases in which sites are found, we can divide the number of times that "yes" results occur *(E)* by the total number of trials *(N)*; the result is the *relative frequency* (in *N* trials) of sites per areal unit. This fraction is often denoted in statistical texts as *r.f. (E)*.

If we were to conduct a random experiment and record the results of each trial, we would see that *r.f. (E)* varies as the number of trials increases. Table 5.1 is a record of such a series of trials. After the first trial, *r.f. (E)* is either 0 (if *E* does not occur) or 1 (if *E* does occur). After the second trial, *r.f. (E)* can be 0 (if *E* does not occur in either trial), 0.5 (if *E* occurs in one trial), or 1 (if *E* occurs in both trials). This process can be repeated *N* times and graphed as shown in Figure 5.2. Over many trials the graph may look like that shown in Figure 5.3 (constructed from data given in Table 5.2).

Figure 5.3 illustrates a fundamental principle of probability theory. As *N* increases, *r.f. (E)* becomes closer and closer to a certain value, usually called *p*. Thus, when *N* is small, *r.f. (E)* varies widely between 0 and 1, but as *N* increases, *r.f. (E)* converges on *p*. Statisticians refer to this phenomenon by various terms, such as "the statistical regularity of chance phenomenon" or "the stability of relative frequencies" (Derman et al. 1973:13).

Regardless of what this phenomenon is called, it lies at the heart of much of probability theory. Returning to the original question of assigning a probability that a location will contain a site, one can see an immediate application of this principle. If we assume that the relative frequency with which a location will be found to contain a site becomes more and more stable as the number of surveyed

TABLE 5.1.

Record of simple results and r.f. (*E*)

Trial	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Result	E	not E	not E	E	not E	not E	E	E	not E	not E
r.f. (E)	1.0	0.5	0.33	0.5	0.4	0.33	0.43	0.5	0.44	0.4

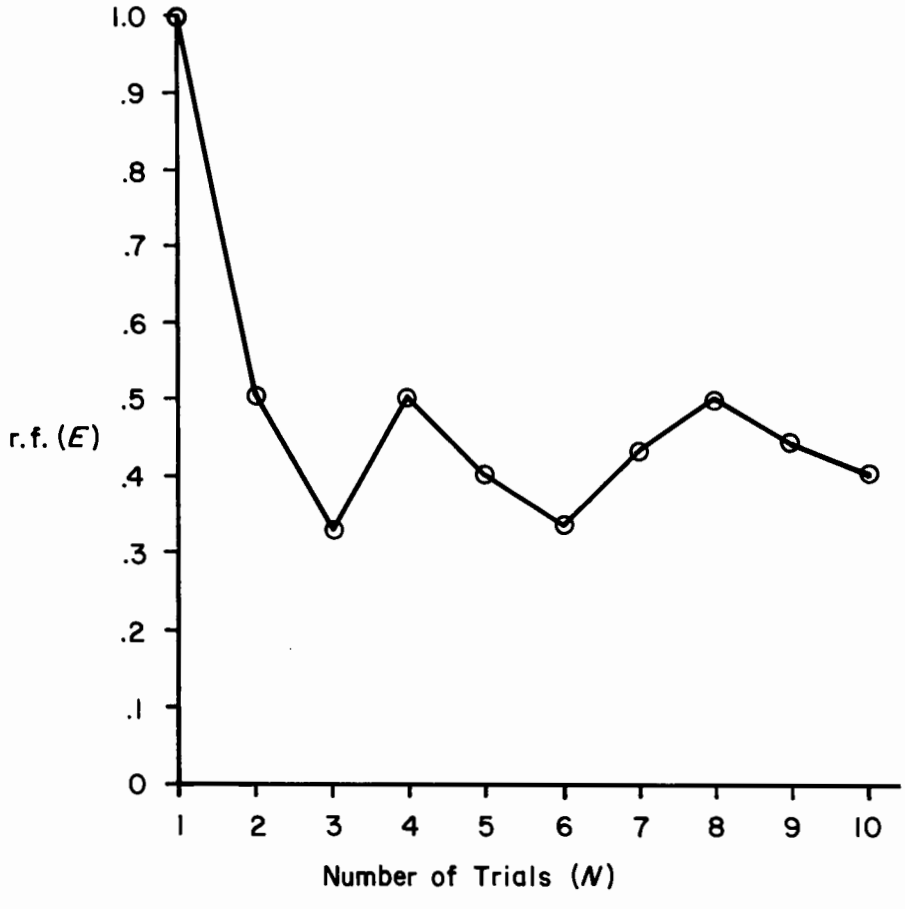


Figure 5.2. Relative frequency of *E*.

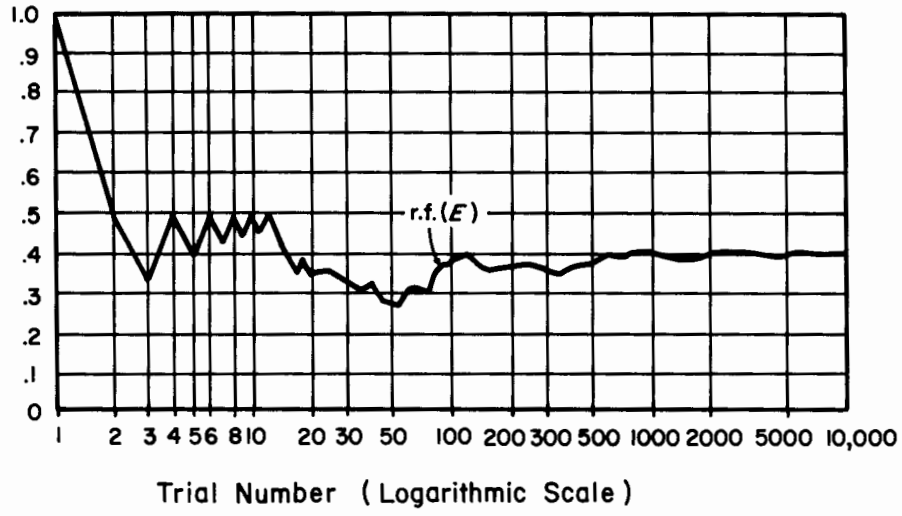


Figure 5.3. Relative frequency of the composite result E.

TABLE 5.2.

Record of how the relative frequency r.f. (E) of the composite result E varies with the number N of trials

Trial	*Frequency*	*Trial*	*Frequency*	*Trial*	*Frequency*	*Trial*	*Frequency*
1	1.0000	30	0.3333	300	0.3567	1700	0.3912
2	0.5000	35	0.3143	350	0.3514	1800	0.3961
3	0.3333	40	0.3250	400	0.3625	1900	0.3974
4	0.5000	45	0.2889	450	0.3689	2000	0.3985
5	0.4000	50	0.2800	500	0.3760	2500	0.4032
6	0.5000	55	0.2727	550	0.3927	3000	0.4003
7	0.4286	60	0.3167	600	0.4000	3500	0.3986
8	0.5000	65	0.3231	650	0.3985	4000	0.3973
9	0.4444	70	0.3143	700	0.3929	4500	0.3953
10	0.5000	75	0.3067	750	0.3947	5000	0.3956
11	0.4545	80	0.3125	800	0.4025	5500	0.3958
12	0.5000	85	0.3529	850	0.4059	6000	0.3978
13	0.4615	90	0.3778	900	0.4078	6500	0.3983
14	0.4286	95	0.3789	950	0.4084	7000	0.3994
15	0.5000	100	0.3900	1000	0.4060	7500	0.4019
16	0.3750	120	0.5000	1100	0.3973	8000	0.4000
17	0.3529	140	0.3786	1200	1.3917	8500	0.4016
18	0.3889	160	0.3625	1300	0.3938	9000	0.4016
19	0.3684	180	0.3667	1400	0.3893	9500	0.3994
20	0.3500	200	0.3700	1500	0.3880	10000	0.4001
25	0.3600	250	0.3680	1600	0.3919		

Derman et al. 1973:Table 3.1

units increases, then it follows that we only need to survey that proportion of the region that is large enough to yield reliable and accurate estimates. Determining how large the proportion should be, how the units should be selected, and a variety of other questions all fall under the rubric of sampling and will be discussed in greater detail in Chapter 6.

The second point that our archaeological example highlights is the practical use of probability statements. How useful is it for a land manager or an archaeologist to know that the probability of finding a site in one unit is 0.02? In most cases the answer is "probably, not very." Instead of being interested in the *absolute* or *marginal* probability of site occurrence, one is usually interested in the probability of site occurrence under specified conditions. For instance, the statements "the probability of site occurrence is 0.0001 in areas with slopes greater than 30°," "the probability of site occurrence in the piñon-juniper zone is 0.15," and "the probability of site occurrence in the piñon-juniper zone and in areas with less than 30° slopes is 0.37" are all much more useful than the general statement that the relative frequency of site occurrence is 0.02.

The probability that one event occurs, based on the information that another or others have occurred, is termed the *conditional probability*. One of the easiest ways to conceptualize conditional probabilities is to use a Venn diagram. In Figure 5.4, the space within the diagram again represents a specific region; the stipled area represents the collection of all site locations. This area is often referred to as the *event set*. It is important to remember that the space within the diagram represents a collection of simple results and does not necessarily imply contiguous land areas.

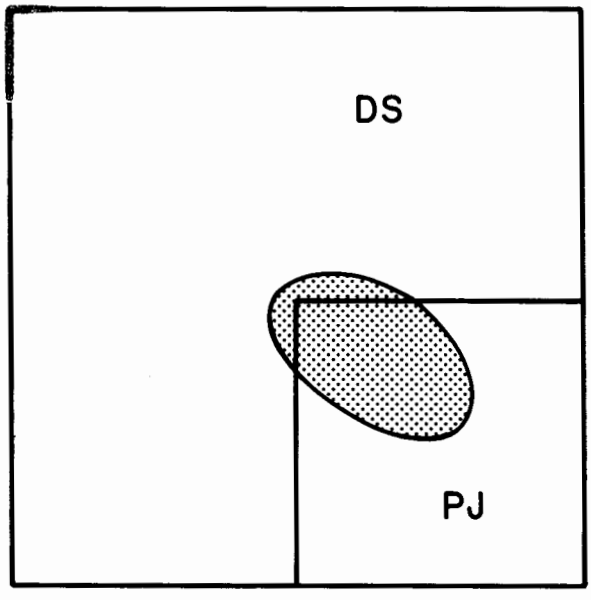


Figure 5.4. Venn diagram showing distribution of sites (shaded area) within piñon-juniper (PJ) and desert shrub (DS) zones.

The region has further been partitioned between piñon-juniper areas, which cover one-fourth of the region when aggregated, and desert shrub areas, which cover the remainder of the space.

Let us assume that the probability of finding a site in a location of specified dimensions in the region is 0.15. That is, the entire stipled area in Figure 5.4 covers 15 percent of the sample space. Most of the stipled area lies in the piñon-juniper partition. If we knew a priori that the survey area was in the piñon-juniper zone, we would want to reassess the probability of finding a site. In the latter situation the sample space would not be the entire diagram but only the partition referring to the piñon-juniper zone (Figure 5.4). Thus, a simple result of finding a site will occur in event set B (a site location) if and only if it is also in event set A (piñon-juniper zone). Stated another way, the outcome can only occur if it belongs to the *intersection* of event sets A and B (denoted $A \cap B$); that is, it will occur only if the surveyed location is in event sets B and A.

The rules of probability calculus are followed to determine conditional probabilities. The conditional probability of B occurring given that A has occurred (designated by $P[B|A]$) is defined by the equation

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

where $P(A \cap B)$ is the probability that both A and B occur. Thus, the conditional probability of B given A equals the probability of both A and B occurring divided by the marginal probability of A.

In our archaeological survey we may have found the following relative frequencies:

	Site Occurrence	
Vegetation	Yes (B)	No
Piñon-juniper (A)	25/400	75/400
Desert shrub	5/400	295/400

Here, $P(A \cap B) = 25/400$, $P(A) = 100/400$, and $P(B|A) = (25/400)/(100/400)$, or 0.25. Thus, if we know a priori that the survey area is in the piñon-juniper zone we can assign a probability of 0.25, *not* 0.15, to finding a site.

In a similar manner, the conditional probability of finding a site in a location given slope, distance to water, or another characteristic could be determined. Further, we could determine the conditional probability of site occurrence given several conditions simultaneously. Indeed, this is what much of predictive modeling is about.

In the following sections of this chapter, complex statistical techniques are introduced that may appear overwhelming to the nonstatistician. These readers should remember that what most of these techniques are trying to do is to partition space in such a way that the conditional probability of finding a site is as close to 1 as

possible for areas likely to contain sites and as close to 0 as possible for areas where sites are probably absent.

Probability theory and calculus lie at the heart of the quantitative aspects of predictive modeling. These subjects have been only briefly discussed in this section. We urge the interested reader and potential predictive modeler to read one or more of the many texts on these subjects (e.g., Derman et al. 1973; Hayes and Winkler 1971; see also Thomas 1976 for anthropological examples).

The Conceptual Model

The approach advocated here for building a predictive model is conservative and comprehensive. The basic approach involves defining several groups, such as sites and nonsites (see Chapter 4 for a discussion of the problems of such definitions) or sites from temporally distinct periods, and selecting a set of independent variables as the potential determinants of site location. For each case (e.g., each recorded site and nonsite location) in each group, measurements are recorded for each of the independent variables, and then some set of mathematical techniques is used to ascertain how different the groups are from each other.

Multivariate statistical techniques are particularly useful in looking at the differences among groups because they simultaneously assess the importance of a large number of variables and can usually be used to calculate the probability of site occurrence given particular values for the independent variables. While the available procedures differ in a number of characteristics, the usual result is a mathematical function or functions that delineate the importance of each variable in defining the groups. If the groups can be separated successfully using these variables, classification functions can be derived that enable us to place cases of unknown group affiliation into the most likely group.

Once the groups have been defined and the classification functions have been derived for the sample locations, the next step is to generalize this information about the probabilities of group membership to the whole population of interest. This is normally done by obtaining data for the same independent variables for locations that were not in the original sample and then using those values in the classification function. In this way a prediction can be made about the probable group membership of each of the measured points in the larger population. Often this is done in a systematic manner that allows the researcher to create a contour or trend surface probability map of the study region.

One problem with the use of these statistical techniques in predictive modeling is that those who use them often do not realize that the use of statistical techniques requires certain assumptions about the data, and that when the assumptions are not met the results of the modeling procedure can be invalid or only approximate. Or, if these limitations are recognized, they are only given scant attention. We argue that the data must be examined using univariate and bivariate

statistics before being subjected to more complex multivariate forms of analysis. It is important that the assumptions associated with a particular predictive modeling technique be identified and that the raw data be evaluated to determine if these assumptions are met.

The first step in building any predictive model should be focused on the variables themselves, especially how they are measured (the scale) and whether this level of measurement is adequate for the modeling technique being considered. Also, it is important to identify the types of distributions that the discrete and continuous random variables possess. Most multivariate techniques are based on interval and ratio scales of measurement (described below) and assume that variables follow a normal, or Gaussian, distribution. The next section of this chapter reviews some of the different probability distributions that a researcher can use to make probabilistic statements about the values of particular variables. Once the distributions of the variables are known and departures from normality assessed, the researcher can decide whether to use normalizing transformations on the variables or to pursue alternative analytical strategies.

The next step in model building is to examine the relationships between pairs of variables using different bivariate measures of association. Bivariate analyses allow the researcher to evaluate the covariance of individual predictor variables and to provide the foundation on which more advanced multivariate procedures are based. If there are problems with the data on a univariate or bivariate level of analysis, there will be problems with more complex analyses. If the steps of univariate, bivariate, and multivariate analyses are all made part of the predictive modeling process, not only will the soundness of the final predictive model be increased, but any weaknesses present in the model probably will have been identified.

Finally, the topic of sampling procedures must be considered. Most predictive modeling techniques are based on the assumption of random sampling, and it is this assumption that permits probabilistic generalizations to be made. Estimates of population parameters are computed differently for different types of sampling procedures, and researchers often fail to take this factor into account. This is an important point, for in constructing predictive models we should be concerned not only with the resultant prediction but also with the amount of error associated with this prediction.

VARIABLES AND SCALES

Those who have constructed archaeological predictive models have tended to concentrate on the sophisticated multivariate statistical models rather than on the basic data. This is unfortunate, because the basic data constitute the building blocks of the models, and they should be thoroughly investigated in the initial steps of model construction. The most cogent reason for a thorough evaluation of the data

is that the most frequently used predictive modeling techniques, such as multiple and logistic regression and discriminant function analysis, are *parametric* statistical techniques. That is, they rely on assumptions about the distribution of the variables being analyzed. Single variables usually are assumed to be normally distributed, pairs of variables are assumed to have bivariate normal distributions, and sets of variables are expected to possess multivariate normal distributions. These assumptions have sometimes been ignored by archaeologists developing predictive models. In contrast, *nonparametric* techniques do not require distributional assumptions.

Variables

For our purposes we will define a *variable* as a measurable entity that is free to assume any of a prescribed set of values. The data used in modeling are the measured values themselves. A variable that can theoretically assume any of an infinitely large and uncountable range of values between two given values is a *continuous variable;* a *discrete variable* can assume a finite range of values, i.e., it can have as many values as there are whole numbers. In predictive modeling, elevation, slope, and cardinal orientation in degrees are examples of continuous variables, while site presence (yes or no) in a sample unit is a discrete variable.

Measurement is the process of assigning a class or score to an observed phenomenon according to some set of rules. What is not always clear, however, is that measurement does not consist only of processes involving numbers. Phenomena can also be classified into types or ranked relative to one another. An important aspect of measurement, especially in a management endeavor, is that the observations be made using an operationally defined process that yields reproducible outcomes that are as valid as possible. This is especially crucial in predictive modeling because we may be considering changes through space and time where some concepts, especially nontrivial ones, may not be readily amenable to measurement or for which no established measurement rules exist. For example, how is the efficiency of site location measured? A simple distance measure could be misleading if slope was not taken into consideration, or if the spatial distribution of the resource itself was not quantified, since different levels of energy expenditure/return could be involved. Efficiency would probably need to be a problem-specific measure, with different rules for measurement being developed in each particular application.

The rules we use to assign a name or number to a phenomenon determine the *level of measurement,* with different rules being associated with different levels of measurement. The level of measurement of a variable determines which methods can be used to analyze it and, ultimately, influences the kinds of inferences that can be drawn from studying its distribution. The level of measurement also strongly influences the type of map that can be used to portray the variable's spatial structure. Stevens (1946) identified the following four basic levels of measurement: *nominal, ordinal, interval, and ratio.* Each level is more rigorously defined than its predecessor, and each contains more information.

The lowest level in Stevens's scheme is the nominal scale. Values are assigned to distinct categories that label or name the phenomenon. The only requirements are that the categories be inclusive—that is, all objects must belong to a category—and that they be mutually exclusive—that is, no object can belong to more than one category. Variables measured on a nominal scale are thus considered to be discrete. For example, if an area is divided into quadrats the archaeologist may be interested in whether each quadrat surveyed contains or does not contain a site. Each quadrat possesses one of these properties, but not both. The nominal scale makes no assumptions about the ordering of or distance between the categories. A nominal scale places limitations on how the variable can be used in statistical operations and cartographic manipulations. In certain situations, however, the values can be counted to form frequency distributions and, if they are spatially referenced, mathematical operations can be performed on their coordinates.

An ordinal level of measurement exists when there is an implied relationship between the classes and they can be ranked (ordered) consistently according to some criterion. Ordinal scales are *asymmetric* and *transitive*. By asymmetric we mean if category A is greater than category B, then B cannot be greater than A. By transitive we mean that if A is greater than B and B is greater than C, then A is greater than C. Variables measured on an ordinal scale are considered to be discrete. In conducting a hypothetical survey, assume the density of sagebrush in 100 by 100 m quadrats was recorded on an ordinal scale using the following five categories: 1 (none), 2 (a few plants), 3 (moderate coverage), 4 (dense coverage), and 5 (total coverage with almost no surface visibility). The asymmetric and transitive characteristics of the ordinal scale can be illustrated using the sagebrush cover ranking. For example, the ranking 2 indicates greater coverage than ranking 1, and 1 can never indicate greater coverage than 2; thus the scale is asymmetric. The scale is transitive because ranking 2 indicates greater coverage than 1, 3 indicates a greater relative cover density than 2, therefore ranking 3 also is greater than ranking 1.

If the categories are ordered and the distances between them are defined using fixed and equal units, the level of measurement is interval. The interval scale lacks a zero point; it can be used to measure differences, therefore, but not absolute magnitude. For example, 80°F is not twice as hot as 40°F because 0 on the Fahrenheit scale is an arbitrary point. To take an archaeological dating example, it would be absurd to say that AD 975 is twice as old as 1950 (Thomas 1976:27). Only when zero points are established by the phenomena themselves can comparisons such as "twice as" have any meaning. Variables measured on an interval scale are considered to be continuous.

A ratio scale has the properties of an interval scale but also possesses an inherent zero point. The defining property of such scales is that any two measurements bear the same ratio to each other irrespective of the unit of measurement. For example, if the distance from point A to point B is 1 mi or approximately 1609.34 m, and the distance from B to point C is 2 mi or about 3218.68 m, the ratio of AB to BC distances is 1:2 in both cases. Interval and ratio data are usually treated together because they frequently can be manipulated arithmetically and statistically in the same ways.

Although data may have been collected at one level, it is possible and often convenient to convert them into a lower level for analysis or graphic presentation. What is generally not permitted, however, is to treat data collected at a lower scale as if they had been measured at a higher scale. For example, it is easy to convert interval and ratio data into ordinal data, but it is not advisable to sum ordinal scores. If one had obtained cover values for sagebrush in terms of area per square meter, they could easily be converted to an ordinal scale of measurement by establishing cutpoints on the original scale. For example, zero coverage could be assigned a ranking of 1, values lying between 0.1 and 0.2 could be assigned a rank of 2, values of 0.3 could be given a ranking of 3, values of 0.4 to 0.7 could be assigned rank 4, and 0.8 to 1.0 could be given rank 5.

In addition to the property of the level of measurement, most variables also have the property of dimensionality. In an archaeological predictive model we want the independent predictor variables to be measures of dimensions that are theoretically related to site location, but not to one another (in a correlation sense). Ideally, a variable is a measure of only one theoretical dimension; when a model includes more than one variable, each one should represent a different theoretical dimension. For example, access to resources may be one dimension in a model, and variables measuring this dimension might include arable land, vegetation types, and elevation. The distribution of vegetation types, however, can be correlated with elevation, and when this is the case in a particular region, it would not be advisable to include both variables in the analysis.

Hybrid variables, representing linear combinations of several other variables, can also be effective predictors if they are not correlated with other independent variables in the same model. If they are uncorrelated with other predictors a hybrid variable still only represents one dimension of variability in a geometric-statistical sense. A basic problem in predictive modeling, and in most aspects of the social sciences, is that one cannot measure a theoretical dimension directly. We can never be completely sure, therefore, that a variable designed to measure one phenomenon is not also measuring part of another dimension at the same time or, for that matter, measuring nothing at all.

When an analysis contains variables that measure the same dimension, there is the possibility that the independent predictor variables will be significantly intercorrelated (multicollinearity) and a statistical model will be produced that has little predictive value because too many variables are correlated with each other instead of with site location. Careful attention to variable selection is a must in predictive modeling, and a shotgun approach, where as many variables as possible are used in the hope that patterns will appear, should be avoided. If it is suspected that there are correlations among some of the variables, such statistical techniques as principal components or factor analysis can be used to reduce the information contained in many variables to a few composite variables, or some of the original variables can be deleted from the analysis. These techniques, which should be fully understood before they are used, are discussed later in this chapter.

Finally, the variables that we choose to represent the theoretical dimensions in the model must be measured using relevant scales of measurement. For example, distances to important resources might be measured in meters and areas of arable land might be measured in hectares or square kilometers. The scale of measurement chosen for a variable will be based on several factors, including theoretical considerations and the precision with which the variable realistically can be measured.

An important class of variables is nondimensional and has values independent of the units of measurement involved. The term nondimensional is used here to mean unitless. It does *not* mean that the variable is measuring an unimportant or non-theoretical dimension of variability in the data. This is an important distinction to maintain. For example, the variable distance to water is often measured on an interval scale in meters. If the values of this variable were converted to z-scores, or standard normal variates, then the variable would be nondimensional. Nondimensional variables are particularly useful in comparisons between sets of variables or in scaling modeling experiments. For instance, it might be extremely difficult to compare the variables distance to water measured in meters (whose scores may range from 0 to the tens of thousands) with slope measured in percent grade (whose scores vary between 0 and 100). Such a comparison could be facilitated, however, by converting both variables' original scores to "nondimensional" z-scores.

Most of the variables currently being used in predictive models can be read from existing maps, such as USGS 7.5- or 15-minute quadrangles, or can be retrieved in machine-readable form from some type of geographic information system. Map-based variables are convenient because the investigator can measure variables for points that were not visited in the field. While this is an important consideration, it also implicitly assumes that the level of resolution of the map from which the information is extracted is sufficient to distinguish critical states of variables. The validity of this assumption should be evaluated in each case in light of what is expected from the model. For predictions of a general nature this assumption may be justified, but as finer and finer predictions are attempted (e.g., "this location will [or will not] contain a site") the quality of the environmental data upon which the prediction rests should itself become the subject of investigation.

Types of Distributions

Within a predictive model, values of variables have specific distributions. These distributions are produced by rules that assign a numerical value to each outcome of an experiment. Where each outcome of an experiment is represented by exactly one numerical value, the rules are called *random variables* (also called *chance* or *stochastic* variables). Because of the difference between discrete and continuous variables, two different types of mathematical models are necessary to describe and analyze the random variables. Discrete random variables are described by the probability mass function and continuous random variables are described by the probability density function.

Probability Mass Function

The probability mass function (Derman et al. 1973) of a discrete random variable has the same information about the probability model for X as does a table of probabilities of the simple events $X = x_1$, $X = x_2$, and so on, where $x_1, x_2, \ldots$ are the possible values or outcomes of X. We denote the probability mass function for X by $p_X(t)$; if no confusion with probability mass functions of other random variables will occur, the x subscript is dropped and we write $p(t)$. The probability mass function of X assigns to every number the probability of the event $X = t$. For each value of X ($x_1, x_2, \ldots$), the probability mass function $p_X(t)$ evaluated at $t = x_1, x_2, x_3, \ldots$ equals a positive number; for all numbers t that cannot be assumed by X, $p_X(t) = 0$. The sum of the nonzero probability values of $p(t)$ is 1.0. The mass function thus gives us the same information provided by a table of probabilities of simple events. The function can be represented on a bar graph that displays probability on the y axis while the x axis is used to represent $t = x_1, x_2, \ldots$, etc. (Figure 5.5a). The probability mass function is useful as a way of quickly gaining a meaningful idea about the probability characteristics of a discrete random variable.

Probability Density Function

For continuous random variables, the function analogous to the probability mass function is the probability density function. To define the density function of a random variable X we use the symbol $f_X(t)$, or when no confusion would result, $f(t)$. The density function of a random variable X may be given a graphical representation, such as the curve shown in Figure 5.5b. In this instance the area under the graph between the numbers a and b represents the probability of an event ($a \leq x \leq b$). The area under the graph of $f(t)$ over the entire horizontal axis is always equal to 1, since the probability is 1 that x is equal to some real number. From a density function, some qualitative conclusions about the variation of the random variable X over repeated independent and identical trials can be drawn. If the density function is nonzero over a line segment from x_1 to x_2 and zero elsewhere (Figure 5.6), then no values outside the range x_1 to x_2 may occur. If the density function is constant over the interval from x_1 to x_2, then all subintervals of equal length (I_1 and I_2) are equally likely to occur. To take yet another example, if the density function is such that most of the area beneath the graph is concentrated in a very narrow range, then repeated experiments on X tend to yield values of the random variable X mostly within the range of numbers where the area is concentrated (Figure 5.7). These simple geometric arguments for calculating the exact probability that a continuous random variable X lies in the interval of numbers between and including a and b are rarely applicable, since density functions normally do not come in the form of rectangles or triangles. To obtain exact answers one must use integral calculus, which provides techniques and formulas for finding areas under curves. These calculations can be quite complex, but tables of probabilities have already been calculated for the types of distributions that most archaeologists will need to consider.

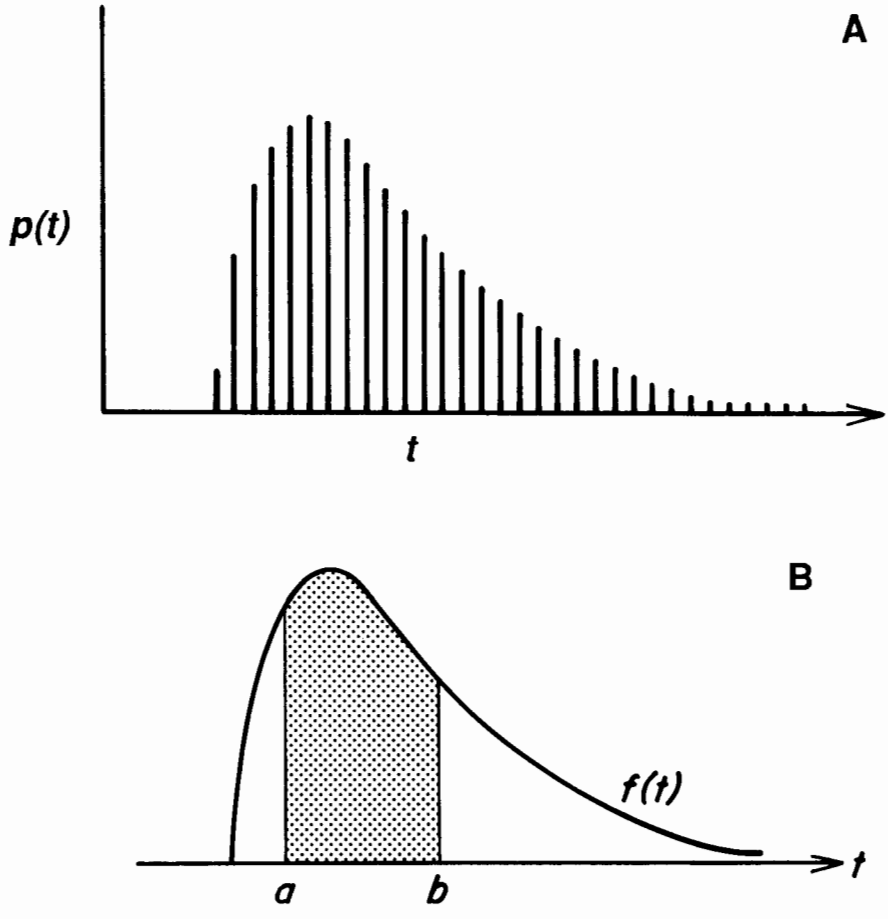


Figure 5.5. Two probability functions. (A) A mass function. (B) A density function. The shaded area is equal to the probability that X lies between a and b.

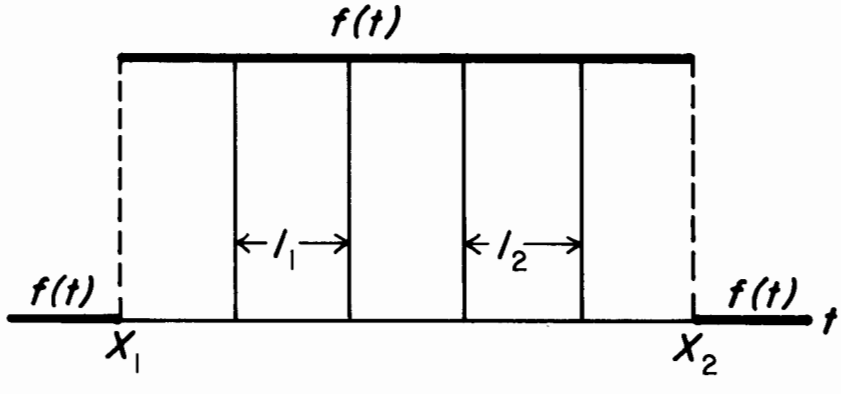


Figure 5.6. Constant density between X_1 and X_2 implies equal probability for the intervals I_1 and I_2, both of equal length.

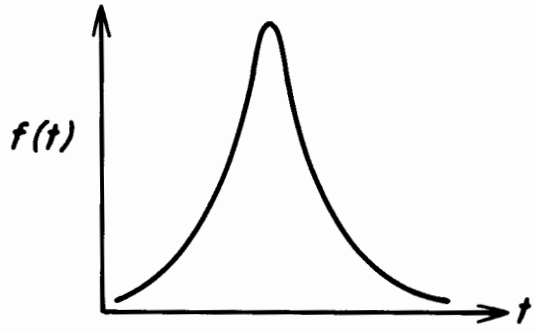


Figure 5.7. A highly centralized density function.

Descriptive Properties of Distributions

For many purposes it is either unnecessary or impossible to obtain all the information contained in the distribution; rather, several descriptive properties that summarize the most important aspects of a distribution may or must suffice. Two of the most frequently used types of information about a distribution are its location and its dispersion. If two density functions have the same graphical shape but concentrate at two different points (x_0 and x_1) on the x axis, the relocation of the graph from x_0 to x_1 represents the only difference between the distributions (Figure 5.8). A descriptive measure that changes values whenever the distribution changes location is a measure of the location of the distribution. On the other hand, a measure of dispersion, or variation, describes how strongly a distribution concentrates about a central value. The measure of dispersion is large when the spread of variates about a central value is large, and it is small when the spread is negligible, becoming zero when all of the probability is at a single point.

The measure of location of a random variable X is commonly referred to as the *mean* of the distribution. Other common measures described in most statistics texts are the *median*, *mode*, and various *quartiles*. The most common measure of dispersion is the *variance* or its square root, the *standard deviation*. The mean of a random variable X is sometimes referred to as the *first moment* of a distribution and the variance as the *second moment*. A generalization of this concept leads to the expected value of the random variable $(X - c)^r$, $r = 1, 2, 3, \ldots$, which is called the rth moment about the point c of the distribution of X. When c is the mean the moments are called central moments.

The other two moments that we will be concerned with are *skewness* and *kurtosis*. Skewness measures the asymmetry of a distribution, and kurtosis provides a measure of how peaked it is. The importance of these moments is that they play a

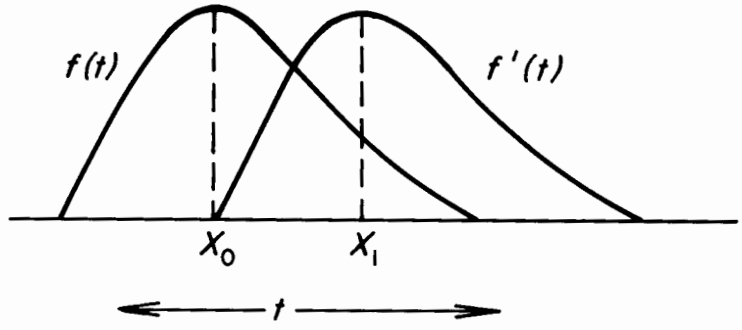


Figure 5.8. Two densities differing in location.

role in approximating the distribution of a random variable. Sometimes the moments of a distribution are known but the distribution itself is unknown, and mathematical techniques have to be used to identify the distribution that possesses those moments. After the distribution has been identified we can then calculate the probabilities of events of interest to us. Formulas for calculating all of these moments are commonly available in basic statistics texts.

There are certain probability distributions that arise quite frequently in many different contexts. Below we will briefly describe some discrete and continuous distributions that are frequently encountered in predictive modeling.

Discrete Distributions

When a discrete random variable has two possible outcomes we have a ***Bernoulli trial*** (Derman et al. 1973). If, for instance, a site can be present or absent, a student can pass or fail, or a stock can go up or down, there are two possible outcomes. One is frequently called a success, the other a failure; the assignment of these terms for possible outcomes is arbitrary. We can define a random variable associated with every Bernoulli trial as follows: if the outcome w is a success, $X(w) = 1$; if the outcome w is a failure, $X(w) = 0$. If p is the probability of success, then $1\text{-}p$ is the probability of failure. A random variable having this probability mass function for some probability p is said to be a Bernoulli random variable, and the resulting distribution is a Bernoulli distribution. The mass function varies with the changing values of p, the parameter of the distribution. If p can be determined for the distribution, this distribution is said to be completely specified.

When we are interested in n independent trials of a random experiment that gives a Bernoulli random variable and distribution, we consider the random variable Z that records the number of successes in n trials. The random variable can assume any of the values $0, 1, 2, 3, \ldots, n$. If $n = 3$ there will be eight outcomes: (SSS), (SSF), (SFS), (FSS), (FFS), (FSF), (SFF), and (FFF), where S denotes a success and F a failure. Corresponding to these outcomes are the following values of Z: 3, 2, 2, 2, 1,

1, 1, 0. To find the appropriate probability model for Z, we find probabilities for the simple events of the composite random experiment made up of the three trials, where each trial is coded as an S (success) or an F (failure). The probability mass function that can be used with the parameters p and n (where p is the probability of success on each of n repeated Bernoulli trials) is termed the binomial and is given as

$$p_Z(k) = \binom{n}{k} p_k (1-p)^{n-k}$$

This equation is used to calculate the number of combinations of n objects taken k at a time. This equals $n!/[(n-k)!k!]$, where ! denotes *factorial*. If n is a positive integer, then the product of the integers from 1 to n is called "n factorial" and is denoted by $n!$. For example, $4! = 4 \times 3 \times 2 \times 1$. Knowledge of both n and p determines the probabilities given in the above equation.

The Bernoulli distribution is a special case of the binomial distribution where $n = 1$. The binomial distribution assumes a fixed number of trials, with the probability of success being the same for each trial and all trials being independent of and not affected by the outcome of the others. As an example of the binomial distribution, a mapped area is divided into eight quadrats or subregions of area a. If there are n subregions, the probability of any point being in a specified region is $1/n$, or $1/8 = 0.125$ in our example. Conversely, the probability of that point not being in the specified region is $q = 1-p$, or 0.875. Using the equation shown above we can calculate the probabilities of there being k points in a quadrat given a total of n points and a probability of p for any one point being allocated to a quadrat. For example, say we wanted to know the probability of there being three points in a quadrat given a total of six points and a probability of 0.125. The probability will be $p(3) = [6!/(6-3)!3!] \times 0.125^3 \times 0.875^3 = 0.16$. In a similar manner we can calculate the probabilities associated with any other number of points.

In many instances these binomial probabilities are not used because they are laborious to calculate and, for most applications, the Poisson distribution described below gives a more readily calculated approximation of the probabilities obtained in an independent random process. The Poisson distribution can be used to approximate binomial probabilities when, considering the parameters n and p of the binomial distribution, n is "large" and p is "small." This works quite well even for modest values of n, say as small as 20 or 30.

Two distributions, the geometric and Pascal, that can be obtained via the binomial distribution are discussed briefly below. The relationships between the geometric, Pascal, and binomial distributions are described in most intermediate to advanced probability textbooks (e.g., Harris 1966). These different distributions are simply used to answer different types of questions.

When the conditions of the independent binomial trials are satisfied, but when one is interested in the *number* of trials required to obtain the first success instead of the number of successes in n trials, the geometric distribution is required. The probability mass function for this discrete random variable is

$$p(x) = (1-p)^{x-1}p$$

For the first success to appear on the xth trial with a probability of p, there must first be x-1 failures, each with probability 1-p. For example, based on prior knowledge let us assume that the probability of a site being present in a quadrat is 0.125. If the site being present is a success then the probability of a failure is 1-0.125, or 0.875, as in the previous example. If one were interested in the probability that the first site would be found in the fifth quadrat surveyed, the formula above would be employed to yield $p(x) = 0.875^4 \times 0.125$, or 0.0733.

The Pascal distribution is also based on the condition of the independent trials being satisfied, but unlike the geometric or binomial distribution, the interest is in the number of trials required to obtain a given number of successes (r). The probability mass function for this discrete random variable is given as

$$p(x) = \binom{x-1}{r-1} p^r (1-p)^{x-r}$$

because if the rth success occurs on the xth trial with probability p, then there must be r-1 successes in the first x-1 trials. This probability function is thus the product of the binomial distribution for r-1 successes in x-1 trials and the probability for success on the xth trial, p. When r is 1, the formula reduces to that of the geometric distribution.

A sample calculation with the Pascal distribution might make the formula presented above less formidable. Assume that from prior research the probability of finding a site (a success) in any one quadrat in a survey area is 0.5. A manager is interested in the probability that the fifth site located will appear by the time the tenth quadrat is surveyed, or in the language of the previous paragraph, that the fifth success will appear on the tenth trial. Values to be used in the formula for the Pascal distribution include $r = 5$, $x = 10$, and $p = 0.5$. The combination notation reduces to the number of combinations of 9 (x-1) objects taken 4 (r-1) at a time, or $9!/5!4! = 362{,}880/2{,}880 = 126$. The remainder of the formula is $0.5^5 \times 0.5^5$, or 0.0010. Then, the probability that the fifth site located will occur by the tenth quadrat surveyed will be $126 \times 0.0010 = 0.126$.

Closely related to a binomial random variable, a Poisson random variable represents the number of occurrences of some outcome, not in a given number of trials but in an interval of time or an area of space. The wide variety of random phenomena giving rise to random variables having this distribution is astonishing. It has been used in control engineering, agriculture, biology, and medicine, to name but a few areas of study. A recent archaeological application is presented by Rogge and Lincoln (1984).

The utility of the Poisson distribution can be demonstrated with the same example used above for the binomial distribution. In many practical applications the quadrat is a relatively small area, implying that k is large and p is therefore small. The probability that a quadrat includes a point may be small, but the number of points n is usually large so that the product Np, the expectation of finding one point

in one area, is relatively constant. If this expectation is called *lambda* (λ), the Poisson distribution is given as

$$p(x) = \frac{e^{-\lambda}\,\lambda^{k}}{x!}\ , \quad k = 0, 1, 2, \ldots$$

where e is the natural constant approximately equal to the number 2.71828. For the present example, $\lambda = 6/8 = 0.75$ since there were six sites in eight quadrats. The quantity $e(-\lambda)$ would be $2.71828^{-0.75} = 0.4723$. Using these values in the Poisson formula, the predicted probabilities for an independent random process where $n = 6$, $k = 8$, and $p = 0.125$ can be calculated. For example, the probabilities of there being three, four, or six points in a quadrat are 0.20, 0.04, and 0.00, respectively.

Probabilities based on the binomial distribution do not apply when we sample without replacement because the probability of a success is not constant from one trial to another. The appropriate probabilities in this situation are based on the hypergeometric distribution, given as

$$\frac{\binom{Np}{k}\ \binom{Nq}{n-k}}{\binom{N}{n}}$$

Assume a random sample of size n is drawn without replacement from a population of N units and that there are k successes and n-k failures, with p denoting the probability of success and q, or 1-p, denoting failure. Let's say a prior inventory of part of a region showed 50 percent of the quadrats contained a site. The remainder of the area, some 300 quadrats, was not surveyed. An archaeologist samples five of the quadrats and finds that two contain sites and three do not. Does this result follow from what was already known about the region? The equation given above can be followed using $N = 300$, $n = 5$, $p = 0.50$, and $k = 2$ to yield a probability that 0.3146 of the 5 units, or 1.573 units, will contain a site. Thus the observation that two units contain sites fits the theoretical observed frequency reasonably well.

Finally, for the binomial probability law to be valid, all possible outcomes of a probabilistic phenomenon can only be classified as either successes or failures. When there are more than two categories of classification, the multinomial distribution applies. More formally, if a probabilistic phenomenon has k possible outcomes with probabilities $p_1, p_2, p_3, \ldots, p_k$, if the probabilities are constant for every trial, and if all trials are independent, the multinomial distribution gives the probability of x_1 outcomes of the first kind, x_2 outcomes of the second kind, through x_k outcomes of the kth kind in n trials. The multinomial distribution could be used in a situation where there were three specific types of sites instead of just a site (or type) that could be present or absent. Blalock (1972:171) notes that a difficulty with the use of the multinomial distribution is the problem of unambiguously specifying a set of outcomes more unusual than the one obtained.

We have mentioned some discrete distributions that frequently arise. When one of these distributions is chosen to model a random variable for which we have observations, the decision generally has either a theoretical or an empirical basis. That is, either the selected probability distribution is a logical consequence of the properties of the phenomenon that are already known to the archaeologist or the probabilities derived from the distribution correspond to the relative frequencies obtained from repeated observations of the phenomenon. In some situations a particular probability distribution may be chosen for both reasons.

Many of the distributions outlined above assume that outcomes are independent of each other, and tests can be devised to determine whether the observed point patterns correspond to independent random processes. These processes are usually mathematically simple and elegant, and they form a useful starting point for spatial analysis. Most geographical applications of models of independent processes are made in order to reject the null hypothesis of independence and randomness in favor of an alternative that specifies some form of spatial dependence (Unwin 1981:60). In order for a clustering of points to be produced, the probability of any quadrat receiving a point cannot be the same for all quadrats. A number of distributions that incorporate spatial dependence exist. Thomas (1977:20–23) notes that spatial analysts have had considerable success in fitting observed frequency arrays based on clustered distributions.

One such clustered distribution is the negative binomial, where the probability of placement increases lineally with the number of points already in a quadrat, leading fairly directly to a clustered point pattern (Draper and Lawrence 1970:99–101). At any one time then, the probabilities of cells receiving a point are not equal, but are directly related to the existing distribution. The probability that a specified quadrat will contain exactly x points is given by

$$p(x) = \binom{k + x - 1}{x} \left(\frac{p}{1 + p}\right)^x \left(\frac{1}{1 + p}\right)^k \qquad x = 0, 1, 2, \ldots$$

where $\lambda = kp$, $\sigma^2 = kp\ (1 + p)$, therefore $p = \lambda k$.

The probabilities predicted by the above equation only depend on the two parameters lambda (λ) and k, where lambda is the point density and k measures the degree of clustering. The value of k falls between zero and infinity, though as k approaches zero the distribution converges on a clustered logarithmic distribution (Bliss and Fisher 1953), and as k approaches infinity the clustering disappears and the negative binomial tends to the Poisson distribution described above. The variance/mean ratio of the negative binomial distribution is always greater than one regardless of the values of lambda and k. Most typically the values of k are not positive integers, and the negative binomial probabilities can be obtained by solving a density function which is an approximation to the formula given above. The approximation is

$$p(x = 0) = \frac{1}{(1+p)^k}$$

$$p(x > 0) = \frac{(k + x - 1)R}{x} \qquad p(x - 1)$$

where $R = p(1+p)$.

Dacey (1968:51–70) describes how the negative binomial distribution can be deduced from different propositions about how the clustered patterns are generated, though for quadrat analyses the most important processes are termed *generalized* and *compound*. The mathematical characterization of these processes is beyond the scope of the present discussion, but it can be noted that the formula for the negative binomial presented above assumes a generalized process that results from some basic affinity among the points in a cluster. For the assumptions of the generalized distribution to hold, one must be reasonably sure that the points and study area are fairly homogeneous in nature.

Compound processes result from a heterogeneity in the numerical population of points under investigation. For example, let us assume that we were interested in the adoption of a particular architectural construction method among the Anasazi in the southeastern Colorado Plateau region. If the prehistoric population density varied significantly over the study area we might observe a clustering in the "adopters" of the construction method not because of short-distance social contacts among the population, but because there were high population densities in lowland areas and low densities in upland regions. The clustering would thus be the result of the lambda parameter (λ) varying over the area, and not the result of a real contagion process. Because the two sets of assumptions associated with the general and the compound processes can lead to the same predicted frequency distribution, the design of a quadrat sampling procedure must specify which model is more appropriate.

In order to fit the negative binomial distribution to a set of data the parameters lambda (λ) and k must be known. Ideally it would be nice if enough were known about the probability model to specify the parameters from a priori information. Realistically this is almost never the case, and one is forced to use some statistical estimation procedure in lieu of theoretical knowledge. Two of the most common methods for estimating the two parameters of the negative binomial are the method of moments and maximum likelihood estimation. An explanation of these two procedures is left to a text such as Thomas (1977:21–23), which provides a cogent discussion of the relevant material and examples. It should be mentioned, however, that the maximum likelihood approach is fairly complex since it involves maximizing the value of a likelihood function.

An alternative that is appropriate in some spatial modeling situations is the Neyman Type A distribution (Ripley 1981:106–107), which involves random placement of a series of initial points from which other points spread randomly. This

probability model can be derived as a compound model resulting from the mixture of two Poisson processes (Rogers 1974).

Wood (1971) has suggested that the negative binomial and Neyman Type A distributions are applicable to prehistoric settlement processes in which clusters of sites are generated by a "contagion" process. Plant ecologists and geographers have used models of this type to describe patterns in which each occurrence of the phenomenon in question increases the probability of additional occurrences nearby (Cliff and Ord 1973; King 1969:45). Clusters of points are Poisson distributed, and each cluster contains one or more points that follow some distribution. If the distribution of points making up a cluster is logarithmic, the overall distribution is a negative binomial; if the points in a cluster are Poisson distributed, a Neyman Type A distribution is yielded (Cliff and Ord 1973; Hodder and Orton 1976). It is assumed that the clusters are spaced far enough apart that a quadrat will not contain more than one cluster, but this will depend on the spatial dispersion within a cluster, the distance between clusters, and quadrat size.

Both of the independent and dependent processes described above lead to probability distributions. The actual observed distribution patterns can then be compared with those predicted by the model, allowing for an evaluation of the likelihood of the observed distribution. One point to be kept in mind when evaluating probability models in most geographical applications, like the negative binomial and Neyman Type A distributions described above, is that the parameters are estimated from the data. Hence, it is likely that the predictive frequencies from many of these a priori models will reasonably fit the observed data.

Continuous Probability Distributions

The continuous distributions described in this section are the *normal, lognormal, t, exponential, gamma,* and *Weibull* distributions. Like discrete probability distributions, continuous distributions are represented in a parametric form, meaning that the general shape of the distribution is given by mathematical equations in which certain constants are left unspecified. For example, in the Bernoulli distribution previously described the p was left unspecified. In the normal distribution described below the mean and variance are left unspecified. When we know the values of these parameters the probability distribution is completely specified, and the probability of any event can be calculated. Put simply, we can determine graphically and numerically the properties of the distribution.

Many of the statistical procedures discussed later in this chapter assume a normal (Gaussian) or at least a quasi-normal (approximately normal) distribution. The general properties of this well-known distribution—the familiar bell-shaped curve—are that it is symmetric, is asymptotic at both ends, has maximum height at the mean, has areas under the curve that represent probabilities of events, and that the distribution of means of repeated samples will tend to be normally distributed. Unfortunately, many researchers fail to determine (even using the most simple, basic descriptive statistics) how well their data meet the assumptions of normality. This can lead to serious deficiencies at more advanced levels of analysis.

Several techniques can be used to judge whether a set of data possesses a normal distribution. A chi-square goodness-of-fit test is one technique that can be used to test whether a sample is from a normally distributed population; however, the test is not capable of identifying some departures from normality. For example, the sample data might possess noticeable skew but the test would not reject the null hypothesis of no significant difference between the distributions. Alternative tests that can detect nonnormal skewness and kurtosis are described below.

To apply the chi-square test the sample data are grouped into classes to form a frequency distribution and the sample mean and standard deviation are calculated. A normal distribution with these parameters is fitted and expected frequencies for each class are obtained. Snedecor and Cochran (1967:70–72) discuss the relevant computations. The chi-square statistic is computed as the sum of (observed-expected)^2/expected for each class. If the data come from a normal distribution, the observed values from the sample will tend to follow the values expected on the assumption of normality and the computed chi-square is small. If the data come from some other distribution, the observed and expected values in each class will tend to agree poorly and the computed chi-square value becomes large. A large chi-square value causes rejection of the hypothesis of normality. As a test for goodness of fit, the most serious limitation of the chi-square test is the requirement for a large sample. As a rule for using the chi-square distribution, each class interval should have an expected frequency of at least 5. Unless the sample is large, only the most frequent class intervals will retain their integrity. The intervals with small frequencies would have to be combined before computing the test statistic, and in doing this, information is lost. When the sample is very small the chi-square test cannot be used at all.

An alternative goodness-of-fit test is the Kolmogorov-Smirnov (KS) one-sample test. This test approaches the normality question by comparing the observed cumulative frequency distribution of the sample to that expected from the population specified by the null hypothesis. The test statistic obtained is the maximum deviation between the observed and the expected distributions. The specifics of this test are described in Lapin (1978:640–644). The KS test is more efficient than the chi-square test for small samples; that is, for a fixed sample size the KS test is more powerful because it provides a higher probability of rejecting a false null hypothesis. A disadvantage of the KS test is that it does not allow the population parameters—the mean and standard deviation—to be estimated from the sample as in the chi-square test. Instead, the population parameters must be specified in advance.

Tests are also available that allow the researcher to evaluate whether sample data with particular values of skewness and kurtosis could have come from a normal population. Skewness coefficients for random samples of normally distributed populations have a mean of zero and a standard deviation of $(6/n)^{1/2}$. The skewness coefficient for sample data can be compared with this value, or two or three times this value, depending on the chosen significance level. For example, if the skewness coefficient exceeds twice this value (+ or -), the null hypothesis of a normal

distribution is rejected at the 0.05 level of significance. This test for skewness is accurate enough when the sample size is greater than 150. For smaller sample sizes, the one-tailed 5 percent and 1 percent significance levels, computed from more accurate approximations, are presented by Snedecor and Cochran (1978:552, Table A6).

Kurtosis can be tested for departures from normality in a manner similar to skewness. Normal distributions possess a kurtosis of 3. When kurtosis is computed, however, the value of 3 is frequently subtracted during the computation so that distributions with positive kurtosis are peaked and those with negative kurtosis are flattened. When computed this way and when very large samples are from a normal distribution, kurtosis is normally distributed with a mean of 0 and a standard deviation of $(24/n)^{1/2}$. Sample kurtosis values that exceed the standard deviation, or exceed it by two or three times depending on the significance level, will lead to rejection of the null hypothesis that the sample data are from a normal distribution. Unfortunately, the distribution of kurtosis does not approach the normal distribution closely until the sample size exceeds 1000. For sample sizes between 200 and 1000 more accurate approximations of the 5 percent and 1 percent significance levels are presented by Snedecor and Cochran (1978:552, Table A7). For small sample sizes, tables of significance levels of kurtosis are not readily available.

The use of the concept of a normal distribution to describe many random phenomena can be theoretically justified by assuming that these phenomena arise from the summation of many statistically independent and identically distributed random causes. We can theorize that a random variable is approximately normally distributed if we can conceive of it as being equal to the sum of a large number of independent realizations of the same random variable X.

The lognormal distribution is the product of many independent realizations of random variables with approximately equal distributions. A nonnegative random variable Z is said to have a lognormal distribution when $Y = \log Z$ has a normal distribution.

Another continuous distribution that is in many respects similar to the normal distribution is the t-distribution. This distribution is also symmetric and has maximum height at the mean, but its shape depends on a parameter called *degrees of freedom* that is largely related to sample size. The t-distribution is commonly used in the construction of confidence intervals related to small samples.

The problems of small samples (those with fewer than ca. 90 cases) and of distributions that are continuous but nonnormal have received little attention in predictive modeling. A fair amount of applied statistical research in disciplines other than anthropology has been directed toward examining the distribution characteristics of small samples from a variety of distributions (e.g., Wallis et al. 1974). Many of these investigations have been instigated by the stark realization that the distributions of variables employed in many models are not quasi-normal. Since an argument can be made from both theoretical and data-oriented perspectives that many archaeological and environmental phenomena are not randomly distributed, some of these distributions are described below.

In the simplest terms, an exponential distribution has the general shape of a reversed "J" and approximates a great many populations where the observations involve items whose status changes through time (or space). This distribution has been used to characterize the expected use life of equipment, for example, and the arrival of cars at a toll booth.

When the distribution of a continuous random variable is considerably skewed, the gamma or Weibull distributions may be useful. The gamma distribution can be used to analyze a system in which the proper functioning of a certain component is essential to the proper functioning of the system as a whole. In order to increase reliability (that is, to increase the length of time before failure), the system may be designed with r-1 spare components that can be used if the critical component fails. When that component fails, one of the r-2 other components takes over. This process can continue until all r components fail, at which time the entire system fails. Assuming that the system can fail only if the critical component fails, the lifetime of the entire system is the sum of the lifetimes of the r components (X_1, $X_2, X_3, \ldots, X_r$). If each of the lifetimes has the same exponential distribution and the same parameters, and if all of the lifetimes are statistically independent, then the gamma probability density function is appropriate. Weibull distributions (Derman et al. 1973:378–390) have also been found to provide good probability models for describing the length of life of certain phenomena.

The exponential, gamma, and Weibull distributions are three classes of distributions that have been used by investigators in many fields to find distributions that explain or describe the variation in nonnegative random variables. Examples of such phenomena include the lifetimes of individuals, travel times, and the lifetimes of biological or even social systems. These distributions provide a reasonable fit to the distributions of many of these random variables, but in other cases the fit is not as close as desired or may even be unsatisfactory.

Statistical Implications of Probability Distributions

The types of probability distributions that variables assume partially determine which statistical tests can be used to analyze those variables. Statistical tests can be divided into two general families: classical or *parametric* tests, which are usually applied to data measured on an interval or ratio scale, and distribution-free or *nonparametric* tests, which can be applied to data measured on nominal, ordinal, interval, or ratio scales.

Parametric tests are generally more powerful and more widely applied than nonparametric tests in predictive modeling contexts, but it is important to note that most parametric tests make certain assumptions about the populations from which the samples are drawn. These assumptions may not always be met, and the data should always be examined to determine whether these assumptions apply. If they do not apply, the extent of the violations should be assessed. The most frequent assumption made about the background population is that it is approximately normally distributed. The smaller the sample, the more important it is that this requirement be met.

If the normality or quasi-normality of the background population is questionable, then a distribution-free or nonparametric test might be required. This is especially true if the variables being tested are derived from small samples. Nonparametric tests constitute a large family of techniques that permit one to cope with frequently unrealistic and limiting assumptions. Another advantage of nonparametric statistical tests is that the theory is sometimes easier to follow. They can also facilitate more efficient data-collection procedures if it is expected that the background population is nonnormally distributed. As Thomas (1976:263) notes, however, whatever the virtues of the nonparametric approach, it remains a second-best substitute for tests based on normality theory. These tests also tend to ignore much sample information that is addressed by their parametric counterparts and therefore may be less efficient. Another handicap is that there are so many nonparametric tests to choose from that the researcher must pay more attention to the added question of efficiency, that is, which test is the most powerful in a particular situation. Additionally, in terms of complex modeling situations many of the nonparametric statistics are not as developed or as applicable as their parametric counterparts.

The other basic distinction in statistical analysis classifies procedures as univariate, bivariate, or multivariate. While most predictive modeling situations are multivariate, a strong argument needs to be made for using univariate and bivariate procedures as a necessary and logical precedent to the use of multivariate processes. If researchers devoted more effort to examining the distributional characteristics of the variables and the relationships among variables, they could determine whether the basic assumptions of more advanced tests were being met and decide whether some variables should be rejected or reexpressed before the variables are incorporated into sophisticated models.

STATISTICAL DESCRIPTION AND INFERENCE IN THE MODEL-BUILDING PROCESS

In the process of building a statistical model we attempt to define and measure the characteristics of individual variables and then to examine relationships between pairs and among sets of variables that affect site location. The characteristics of the individual variables and the measures of association are then generalized to a population. A large battery of univariate, bivariate, and multivariate statistics can be employed in the process of generating predictive models of site location. Univariate statistics are generally used to elucidate the various characteristics of probability distributions associated with particular archaeological and environmental variables. Such descriptive statistics as the mean, median, variance or standard deviation, skewness, and kurtosis, coupled with such graphic displays as histograms and cumulative probability function curves, permit us to determine whether the variates are distributed in a somewhat normal fashion. As described above, this assessment of the distribution is critical because normality is a fundamental assump-

tion of most parametric statistical tests. If the variates are nonnormally distributed, statistics and graphic procedures can aid us in determining an appropriate normalizing transformation. Alternatively, one of the other distributions described above may provide a good fit to the observed data.

Exploratory data analysis (Tukey 1977), which emphasizes the use of visual displays, can help to describe the univariate, bivariate, or multivariate distributions of variables. As Hartwig and Dearing (1979) note, the basic philosophy underlying exploratory data analysis is one of searching a data set using a number of alternative techniques in order to maximize what can be learned. They feel that a potential problem may arise when data analysis is equated with statistics, that is, when numerical summaries of the data are used to the exclusion of other methods of analysis.

In contrast to a traditional statistical approach, exploratory data analysis does not impose a hypothesis of some pattern on the data; it lets the pattern emerge from the data. It also emphasizes the reexpression of variables that are not normally distributed or might be expressed better on a different scale. Data can be reexpressed by any transformation as long as the discovered patterns can be related back to the original data.

At the most elementary level, traditional descriptive statistics aid in the detection of outliers, values that by their very label indicate something outside the range of the main body of variates, something that is anomalous. Sometimes these values may lie three, four, or even more standard deviations from the mean. The detection of outliers and decisions on how to deal with them at the beginning stages of quantitative analysis are essential because of the pathological effect that these values can have on the final results. Although rules have been proposed for rejecting outliers, automatic rejection is not always advisable. It is possible that the outlier provides information that other variates cannot because it arises from an unusual combination of circumstances. From a managerial perspective, the identification of outliers is important because they represent something that does not follow the norm and that requires further investigation.

One of the most fundamental concepts of the exploratory approach is a breakdown of the data into *smooth* and *rough* components. The underlying structure of a set of observations is smooth when a straight line depicts the relationship between two variables or a curve depicts the distribution of a single variable. Smoothness represents regularity in the data. When the smooth data are extracted from a data set, what remains are the deviations or residuals, the differences between the smooth and the actual data, which are called the rough data. The most desirable situation is a rough data set that contains no additional patterns or structure.

Exploratory data analysis, like some traditional univariate statistical analyses, places a premium on analysis of single variables in an attempt to understand the central tendency, variability, and shape of the distribution for each variable. Techniques employed include the stem-and-leaf display, box-and-whisker plot,

and resistant number summaries. A stem-and-leaf display is a cross between a rank-ordered list and a histogram. To create a stem-and-leaf display each observed value is separated into its first digit and remaining digits. Each number that occurs one or more times as a first digit in the data set is listed vertically in ascending order, and a vertical line known as the stem is drawn to the right of this column of numbers. The remaining digits are listed horizontally in ascending order from left to right in the same row as the first digit to which they belong, creating leaves. This results in a histogram that retains and ranks all of the observed values while losing none of the data. When the observed values have many digits, rounding the numbers to a few digits may simplify the display. Furthermore, these displays can be stretched or condensed by subdividing each row into two or more divisions or by combining two or more adjacent rows.

This description of the stem-and-leaf diagram is included as an example of how exploratory data analysis attempts to extract information from a data set. On a univariate level, exploratory data analysis identifies and describes major characteristics of distributions using measures of location and spread that have the property of resistance. The term *resistance* means that these measures are not highly sensitive to departures from the normal distribution and thus they are suitable indicators of location and spread for a wide variety of distributions (Hartwig and Dearing 1979:19). The exploratory data analysis approach also uses several summary statistics, rather than just one or two, to summarize information about a distribution. A box-and-whisker plot can be used in addition to numeric summary measures to portray the major characteristics of a distribution. These plots provide detail when it is often needed the most—when the tails of a distribution contain extremely large or small values. Tukey (1977) provides a detailed discussion of these measures and plotting procedures (see also Clark 1982 for a cogent discussion of archaeological applications).

Traditional bivariate statistics include many procedures appropriate to the four different levels of measurement. Because they indicate how one variable is related to another, matrices of bivariate statistics frequently provide a starting point for multivariate procedures that, in turn, evaluate relationships between and among a large number of variables. At the nominal and ordinal levels of measurement we employ nonparametric tests like chi-square, gamma, lambda, Kendall's tau_a and tau_b, rank-order correlation, sign tests, and the like. Parametric bivariate tests include analysis of variance, differences of means, and Pearson's r and are chosen according to the level of measurement. Outliers can also be detected with bivariate statistics and graphic techniques, but if we are using parametric bivariate techniques on nonnormally distributed variables, problems with the reliability of results and predictions will begin to appear.

When examining bivariate relationships, the researcher may be faced with the problem that both variables being considered are strongly affected by a third. One way of examining this problem is by cross-tabulation, which is a joint frequency distribution of cases according to two or more classificatory variables. The joint frequency distributions can be analyzed with a variety of statistics to determine

whether or not the variables are independent. The statistics describe the degree to which the values of one variable predict or vary with those of another. The relationship between two variables can also be calculated while controlling for the effects of a third. Examining a bivariate relationship while controlling for a third variable can frequently be a problem with archaeological data bases that contain a limited number of cases, however, because each addition of a category to each variable in the relationship exerts a drain on the average cell frequencies in a cross-tabulation. A very large sample is needed to generate relatively simple controls.

Partial correlation is another technique that provides a researcher with a measure of association describing the relationship between two variables while adjusting for the effects of one or more additional variables. Conceptually it is analogous to cross-tabulation with control variables, but in this situation the control is statistical rather than literal. Partial correlation is based on the simplifying assumption that the relationships among variables are linear and that the effect of the control variable is linear throughout its range. Once the relationships among the independent, dependent, and control variables are determined, it is possible to predict the values of the dependent variable using the independent variable, while controlling for the influence of the other variable(s). The advantage of including partial correlation as one of the steps in building predictive models is that this procedure can help to detect spurious relationships—that is, those relationships between two variables *A* and *B* that are solely the result of variable *A* varying with some other variable *C*, which may actually be the true predictor of variable *B*. When variable *C* is controlled for, variable *B* may no longer vary with *A*, and the spurious relationship will have been detected.

Exploratory data analysis also provides techniques for looking at the relationships between pairs of variables. Important factors that are considered with this approach include the shape, strength, and direction of the relationship. Scatterplots, traces, and smoothers, such as the Tukey line (similar to a regression line), are used to examine pairs of variables for nonlinearity not removed by reexpression during the univariate stage of analysis. If nonlinearity is apparent and not due to just a few deviant values, some linearizing transformation on the independent and/or dependent variable is required. As in the univariate approach, the analysis does not stop with an examination of the smooth data set. The importance of subjecting the rough data to the same careful examination given to the observed values is emphasized. Hartwig and Dearing (1979) provide convenient summaries of exploratory data analysis approaches to examining bivariate relationships.

The use of techniques such as those outlined above before constructing complex multivariate predictive models is necessary for several reasons. First, these techniques assist us in determining which variables meet the distributional assumptions of the parametric statistical techniques that we prefer to use. If the variables do not meet the criteria for parametric statistics, nonparametric alternatives may be appropriate. Second, these techniques define the relationships between pairs of variables, and third, they make it possible for us to determine, and untangle, the

complex relationships that can exist among several variables. Ultimately, all of these techniques help us to identify a set of independent variables that are related to site location; just as important, they show us how the set of predictor variables are related to one another.

Establishing a relationship between one or more interval- or ratio-level independent variables and one or more interval- or ratio-level dependent variables can range from being a relatively simple procedure to being a complex one. At the simplest level, bivariate regression can be employed, and when more than one independent variable is involved, a multiple regression scheme of one form or another can be used. If the independent variables are intercorrelated, a modification to the normal approach may be required. A principal components analysis can be used to create a new set of orthogonal (uncorrelated) variables that are linear combinations of the original ones (Harris 1975:163–167). Or, in multiple regression situations, ridge regression or latent root regression may represent a viable alternative (Gunst and Mason 1984). When there is a set of dependent variables in addition to the set of independent variables, canonical regression with the original variables or with the principal components of each data set is required (Harris 1975:132–146). Some of these techniques and their underlying assumptions are described more fully later in this chapter.

DEFINING SITE CLASSES

Earlier we defined the process of constructing quantitative predictive models as consisting of group discrimination and classification. The former involves taking two or more predefined groups and producing a mathematical function that describes the use of a set of independent variables to separate these groups. Classification, on the other hand, involves capitalizing on any group differences that might be present in order to develop an algorithm for classifying other elements of the population into the most appropriate group. To be successful at discrimination and classification requires choosing and understanding the distributions of appropriate variables, as we have discussed, and ensuring that the groups we are using are relatively homogeneous. Several problems can affect our ability to form homogeneous groups; of these, we will discuss temporal and functional variability, the definition of site classes using cultural resource management data, and the use of nonsite locations as a group in predictive models.

Temporal and Functional Variability

In areas where the temporal dimension is long enough to incorporate adaptational changes, we must be able to control for these temporal and functional factors before we can make realistic predictions. For modeling purposes, we assume that contemporaneous sites are distributed over the landscape according to various

social, economic, and environmental factors, and that these factors determine the locations and types of sites found. Thus, different site types may be associated with different configurations of environmental and social variables, and these relationships may change through time.

The problem of determining site function is compounded when a region has been occupied for an extended period of time, possibly by more than one cultural tradition involving more than one economic pattern. In these instances the associations between site types and environmental features may change through time. During one period basecamps may be found near floodplains while during another they may be found on ridge crests. In addition to controlling for functionally distinct site types we must also control for temporal variability in sites and possibly in the environment.

The importance of delineating temporal and functional aspects of sites and the environment has long been recognized by the archaeological community. In early settlement-pattern studies this was accomplished by taking large surface collections from each site and relating them to the results of intensive excavations of samples of each site type (see MacNeish 1964 and Sanders et al. 1979 for examples of this approach). In cultural resource management, however, archaeologists are not usually in a position to examine settlement data in this way. Surface collections tend to be small, and test excavations are usually not directed at linking surface materials to subsurface remains. This situation thus calls for a different approach.

Defining Site Types with Cultural Resource Management Data

Archaeologists developing predictive models with cultural resource management data have not always successfully dealt with questions of temporal and functional variability. Often little attention is devoted to this problem, and modelers lump all sites together in an analysis (Grady 1980; Larralde and Chandler 1981). There are two major reasons for the failure to develop usable site classes in predictive modeling. The first is that many predictive models are generated from small samples. Scholtz (1981) developed the Sparta Mine predictive models of prehistoric and historical site locations on the basis of 37 and 31 sites, respectively. If we develop models on the basis of small samples that represent our only cultural resource knowledge of the area, then we may not know enough about the area to evaluate the accuracy and precision of the model. Even if the data are derived from probabilistic sampling techniques, it is important to remember that unless the sampling fraction is large enough, say above 20 percent, we have not sampled a significant proportion of the population (Cowgill 1975). If we have reason to believe that the survey area is similar to surrounding areas, then site location data from those regions can be used to augment the data base when developing a predictive model. As discussed in Chapter 7, the poor use of existing data has hampered the use of predictive modeling in archaeology.

The second reason that site classes are sometimes poorly developed in predictive models is a reflection of a general trend in cultural resource management. The determination of temporal and functional attributes from surface collections of artifacts usually requires fairly large samples. Even when artifacts are numerous it still may be difficult to define site classes, and in the case of minimal artifact scatters, this problem is exacerbated because even if every artifact were collected and analyzed the sample might not be sufficient to yield a temporal or functional designation for the site. Given the growing trend among some federal agencies to limit the number of artifacts collected in the field, archaeologists often are forced to do in-field analyses. It is not surprising that many surveys result in a disproportionate number of undated sites.

In the absence of artifact data that could be used to assess site function, archaeologists have often used the number of artifact types as a measure of occupational intensity, which in turn is used as a proxy for function. One problem in using this approach is that collection and vandalism can seriously skew surface assemblages of sites. Often, therefore, sites with less assemblage diversity than expected have simply been heavily collected. Another problem with this approach, even if a site has not been heavily collected, is related to sample size. If 100 artifacts representing 27 types are recovered at one site, while at another there are 70 artifacts with 20 types represented, can it be said that the first site has a more diverse collection simply because there are more types? And can this be taken one step further to say that the first site represents greater occupational intensity than the second? In many cases we need to know whether the collection from the first site contains more types than we would expect given its sample size. Several procedures have been used in an attempt to answer this question.

The first procedure is applicable when the number of artifact classes represented can be determined from in-field analysis. Kintigh (1984) devised a method for measuring assemblage diversity by simulating the composition of a large number of theoretical samples drawn from a range of total artifact classes for a given sample size. To compare the diversity of two sites one looks not at the assemblages themselves but at the expected diversity for each site sample given its size. With this procedure one can group sites objectively into classes that may have theoretical significance using some form of clustering technique (Everitt 1974). Sites with less diversity than is expected may be limited activity loci, while sites with expected or greater than expected diversity may represent occupational loci.

A second approach to site classification based on cultural resource management data is to use an empirical Bayesian technique (Carter and Rolph 1974; Chernoff 1982; Efron and Morris 1973, 1975, 1977). In this case one uses a pooled estimate of the population's proportion to obtain more reliable estimates for individual sites. This technique has the advantage of retaining information from sites with large collections that provide the most reliable estimates. In cultural resource management situations, artifacts from all sites in a subregion could be used to develop a pooled estimator (weight) that could then be used to recalculate the proportions of various artifact classes in each site in a larger region. These refined

estimates could then be put into a clustering algorithm that could be used to obtain objective site classes.

Heterogeneity of Sites and Nonsites

In many predictive modeling situations archaeologists have adopted the concept of a binary response variable; often locations are classified as either sites or nonsites. This approach may make it difficult to operationalize some discrimination and classification procedures. It may be possible to derive a numerical function that defines how these two groups are separated, but cases in the site group may be widely dispersed around the group centroid (intuitively the "center" of a group of points) because of temporal and functional variability. This degree of dispersion may lead to a high percentage of misclassifications. Additionally, from a managerial perspective, information on different types of sites and their distributions through time may be not only important but required—cultural resource managers would undoubtedly have different strategies for managing small lithic scatters and large Pueblo III sites.

Additional problems can result from the use of a binary response variable when one of the groups represents nonsite locations. In such cases, nonsite locations generally are used as a control group, permitting the researcher to identify patterns in the environmental contexts of sites that form the other group. It is not clear whether the importance placed on nonsites is justified (see Chapters 7 and 8 for an alternative view), and this grouping of all nonsites may cause statistical problems when the nonsite category is heterogeneous. The members of this group may not necessarily have relationships among themselves, yet most classification techniques assume that there is less variability within than between groups. The problem is that we may be trying to distinguish one group (sites) from another group (nonsites) that consists of a random assortment of elements of the population. Essentially, the site group may be a subset of the nonsite group, and cases that actually belong in the site group may be placed in the nonsite group because the latter represents the variation of the entire data set. Only those cases located near the center of the site group may actually be classified as sites. Because the nonsite group may be heterogeneous, a large number of cases may be required in order to represent the dimensions of environmental variability accurately. If the sample size of the nonsite group is small, the chance that the environment will be badly represented is increased, as is the possibility that the centroid of the nonsite group may vary from one analysis to another. Our predictions as to whether a location will be a site or a nonsite may therefore change from one analysis to another.

The results of a recent predictive model for the Fort Carson Military Reservation provide a useful example of these problems (Altschul and Rose 1986). During a 30 percent simple random sample survey of the base, 98 prehistoric sites were recorded (Alexander et al. 1982). In the course of constructing a predictive model we decided to test the notion that sites and nonsites would form relatively homogeneous and distinct groups based on their "environmental" composition. Each 250 by

250 m quadrat that contained a site was scored on eight environmental variables, as were an equal number of randomly selected nonsite quadrats. A Pearson product-moment correlation coefficient, a bivariate technique for normally distributed data, was computed between each case and every other case. The resulting matrix was analyzed through an agglomerative hierarchical clustering algorithm (see below).

Our expectation was that groups of sites, ideally recognizable as site classes, would form and would be statistically differentiated from either one group of nonsites or several groups of nonsites, each representing a separate environmental zone. The results were enlightening. As expected, several groups were easily distinguished. What was not expected, however, was that each group was composed of nearly equal numbers of sites and nonsites. The results suggested that, as a group, sites were not distinguishable from nonsites. The analysis forced us to reexamine our approach, and we concluded that one confounding factor was that Fort Carson itself probably was not a useful analytical unit. The base was then subdivided into three major drainage basins and the analysis was repeated, with much better results.

We do not mean to imply that all predictive models using a site/nonsite binary response variable are inaccurate or lead to invalid predictive models. Indeed, this approach may be dictated by sample size considerations because the statistical characterization of multiple groups requires that each is adequately represented by a sufficient number of entities. We simply suggest that this aspect of predictive modeling needs more critical evaluation and that when sample size is sufficient there is a need for a response variable with multiple categories.

Defining Site Classes and Reducing Heterogeneity

Multivariate parametric techniques, such as cluster analysis and principal components or factor analysis, provide the means to define classes of phenomena (groups) and reduce the amount of variability present in a data set. The technique of cluster analysis is mentioned briefly first because some clustering techniques can be used with variables measured with scales ranging from nominal to ratio. Principal components and factor analyses, on the other hand, require sets of independent variables measured on interval or ratio scales. Cluster analysis will place sites into relatively homogeneous groups; principal components and factor analyses are valuable techniques for handling the problem of multicollinearity within a set of independent predictor variables.

Cluster Analysis

Cluster analysis refers to a set of techniques that can be used to subdivide a data set into constituent groups. Some methods group *cases* on the basis of observations made on a set of variables (Q-mode) while others group *variables* (R-mode). Our interest in this context is in Q-mode analysis. Ideal data for cluster analysis

would yield clusters so obvious that in situations with relatively small numbers of sites the groups could be visually determined. In practice, however, the situation is not so simple, and as a result there has been a proliferation of clustering techniques.

Most clustering techniques begin with the calculation of a matrix of similarities or distances between entities. Similarity coefficients have been widely discussed in the literature (Sokal and Sneath 1963), where they are sometimes called measures of association. A similarity coefficient measures the association between two cases (such as sites) given the values on a set of variables common to both. Values of the coefficients normally range from 0 to 1, with 0 implying no similarity and 1 designating perfect agreement, but values of different similarity coefficients applied to the same data may vary widely in comparison with one another. One of the most common situations in which these coefficients are used is with variables of the binary response, presence/absence type.

A very large number of similarity coefficients have been proposed, primarily because of uncertainties about how negative matches should be incorporated and whether matched pairs of variables should be equally weighted or carry twice the weight of unmatched pairs or, alternatively, whether unmatched pairs should carry twice the weight of matched pairs. That is, if we have two binary response variables coded + and -, a two-way association table has the following four possible cells: + on one variable and + on the other, + on one and - on the other, - on one and + on the other, and finally, - on both variables. This last combination of two negative matches lies at the heart of the problem. Some coefficients exclude negative matches while others give higher weightings to matched pairs. Different similarity coefficients may have very different values on the same set of data for these reasons.

Some coefficients have been devised specifically for use with data measured on interval and ratio scales (e.g., the correlation coefficient), and Gower (1971) has defined a general similarity coefficient that can be used for data measured on any scale. Gower's similarity coefficient can also be used when the data set contains variables measured on different scales.

Distance measures can also be used as the object of cluster analysis. A distance measure is a numerical function $d(x,y)$ of pairs of points of a set. This function is said to be metric for the set if it satisfies the following three conditions:

1. $d(x,y) \geq 0$; $d(x,y) = 0$ if $x = y$
2. $d(x,y) = d(y,x)$
3. $d(x,z) + d(y,z) \geq d(x,y)$

The last condition is the one that separates distance measures from similarity coefficients; it is referred to as the metric inequality or the triangular inequality. Most distance functions can be transformed into similarity measures, but the reverse process is much more difficult because the triangle inequality must be satisfied. One of the most widely used distance measures is the Euclidean distance (Everitt 1974:56), but it may be unsatisfactory when used on raw data since it is strongly affected by the scale of a variable. Other common distance measures

include the absolute or city-block metric (Carmichael and Sneath 1969) and that of Mahalanobis (1936).

Many clustering algorithms are available, but only agglomerative hierarchical clustering and partitioning techniques are discussed here. The basic procedure for most agglomerative hierarchical clustering techniques involves calculating a similarity or distance matrix between entities. The end product is a dendrogram showing the successive fusion of cases, culminating with all the cases in a group. The differences among the various hierarchical clustering methods lie in the procedures used at each particular step to fuse cases or groups of cases that are the most similar or the closest to each other. Different methods use different means of defining the distance or similarity between a case and a group or between two groups.

Partitioning techniques differ from hierarchical techniques in that they permit relocation of entities, thus allowing poor initial partitions to be corrected at a later stage. Most partitioning methods are formulated to partition a set of cases in a way that optimizes some predefined criterion, such as the trace (sum of the elements of the main diagonal) of the pooled within-group sums-of-squares-cross-products (SSCP) matrix (see discussion of discriminant function analysis below). Most of the methods assume that the investigator knows in advance how many groups there are, although some do permit the number to be changed during an analysis. For example, if we had 100 sites and we had reason to suspect that there were five temporally or functionally distinct site types, we could specify that the sites were to be partitioned into five groups. Other methods require an initial specification of the cluster configuration, or what the membership of the clusters will be like, based on prior knowledge.

Many different methods have been proposed for initiating clusters, which is normally the first step in a partitioning type of cluster analysis. Each case is then put in the cluster whose centroid is closest to the location of that case. For example, the *k*-means clustering program in the BMDP package (Dixon 1981) uses the Euclidean distance to measure the distance between each case and the center of each cluster. Relocation of cases to another group takes place in an attempt to optimize some clustering criterion—e.g., to minimize the variability within a group or to maximize the variance between groups. Regardless of the particular clustering technique used, the objective of applying these techniques in predictive modeling is to define clusters of sites that might have temporal and/or functional designations.

Principal Components and Factor Analyses

In many predictive modeling situations, measurements are made on a set of independent interval- and ratio-scale variables for each case, such as a quadrat. Usually these variables measure some aspect of environmental variability. Ideally, each variable represents just one dimension of variability, and this is one of the assumptions of such statistical procedures as multiple regression, logistic regression, and discriminant function analysis. Unfortunately, variables are not always

uncorrelated with one another, meaning that one variable's values may be partially a function of another variable. *Principal components analysis* is a data-transformation technique that can be employed to create a new set of variables that are linear combinations of the original variables (Daultrey 1976; Harris 1975; Morrison 1976). The original data are transformed so that the same amount of variability is described using the same number of variables but in such a way that

1. the first axis (linear combination of the original variables) accounts for as much of the total variance as possible;
2. the second axis accounts for as much of the remaining variance as possible while being uncorrelated with the first; and
3. the third axis accounts for as much of the remaining variance as possible while being uncorrelated with the first two, and so on.

When significant correlations are present among a set of variables, normally a few large axes account for a substantial percentage of the total variance while a larger number of variables account for smaller amounts of variance. The small axes accounting for only small amounts of variance are normally discarded from further analysis. Thus, the investigator has transformed an initial data set of p correlated variables into a data set of m uncorrelated variables that explain most of the variance, with m normally being much smaller than p.

The creation of this new set of variables, or principal components, has several advantages. The first is that the variables are not correlated with one another—that is, each one measures a separate dimension of variability. This is one way of meeting the predictive modeling assumption that no significant linear relationships exist among the independent variables. The second advantage of principal components is that a large amount of variance in the original data set is explained by a smaller set of variables, introducing a parsimony that is normally desirable in any scientific analysis. By examining the relationships between the original variables and the principal components it is frequently possible to interpret the meaning of the principal components in terms of the original variables. The focus of interest when principal components analysis is used as a data transformation technique, however, is on the scores exhibited by the individual cases on the principal components. Each case, such as a site, will have a score on each of the principal components defining some aspect of variability among the original variables. These scores can be used in subsequent statistical analyses in lieu of the values for the original variables.

As many principal components are needed as there are variables in order to reproduce the intercorrelations among all of the original variables. If the principal components accounting for relatively small amounts of variance are eliminated, a more parsimonious description of the original data has been obtained, but it has been obtained at the expense of possibly losing the ability to reproduce the intercorrelations among the original variables. It should also be noted that principal components analysis makes use of all the information about every variable, though it may be that some of the variation in a case's scores on a given variable is unique and attributable to things that have nothing to do with other variables in the set.

When this unique variance is eliminated from the analysis we might be able to provide a better explanation of the relationships among the variables.

In principal components analysis the new linear combinations of variables that are produced are uncorrelated with one another, and each successive principal component accounts for less variance than its predecessors. If the investigator suspects that the true factors determining the structure of the data are all of about equal importance, then the technique of factor analysis may be more appropriate than principal components analysis. It must be mentioned, however, that many authors regard principal components analysis as a form of factor analysis and frequently use it as a first step in such a study.

The term *factor analysis* refers to a family of techniques that correct for one or more of the shortcomings of principal components analysis. Common to all factor models is the explicit separation of unique variance from variance held in common among variables and the assumption that the observed correlations among variables are generated by a smaller set of "latent" variables. Depending on one's preconceptions about the nature of the underlying variables, each variable's communality (percent variance held in common with other variables) may have to be specified in advance. By employing factor analysis instead of principal components analysis, a researcher gains the ability to reproduce the original pattern of intercorrelations among variables from a relatively small number of factors. What is lost is the straightforward relationship of a case's scores on the original variables and its scores on the various factors.

Another loss in almost all forms of factor analysis is the uniqueness of the solution. A given factor structure simply represents a description of the original intercorrelations. Unless additional constraints are imposed, the correlation pattern can be described by any other frame of reference employing the same number of factors (Harris 1975:26). Most factor analysis methods employ some type of arbitrary constraint to obtain a preliminary factor structure and then rotate the frame of reference until a factor solution is found that comes close to some prespecified set of criteria. In many predictive modeling situations in which we simply desire a straightforward transformation of the data into a new set of uncorrelated variables, principal components analysis adequately accomplishes this task. More sophisticated types of factor analysis are usually appropriate when a researcher is interested in obtaining a better explanation of the relationships among a set of variables. Whatever the case, factor analysis is a complex form of multivariate statistics that should be used cautiously and with understanding.

MODELING TECHNIQUES

Predictive modeling takes place in steps, and we have presented a number of steps that should precede the use of complex multivariate modeling techniques. We have emphasized that the researcher should choose variables that are theoretically relevant and that represent different dimensions of the model. Most modeling

techniques make certain assumptions about how the variables are measured and how they are distributed. Thus, the steps using univariate, bivariate, and multivariate techniques to understand variable distributions and relationships are important, especially if some variables need to be reexpressed to meet the assumptions of the multivariate techniques. In addition, groups must be carefully defined so that subsequent multivariate manipulations produce realistic classifications. Three parametric multivariate modeling techniques—general linear regression, logistic regression, and discriminant function analysis, are described below.

General Linear Regression

Regression models are frequently used in predictive modeling situations when there is a dependence relationship between the dependent variable and one or more independent variables. If it is assumed that a relationship exists between a dependent variable Y and k explanatory variables $X_1, X_2, X_3, \ldots X_k$, the relationship can be expressed as

$$\hat{Y}_i = \hat{b}_0 + \hat{b}_1 X_{i1} + \ldots + \hat{b}_k X_{ik} + e_i$$

for $i = 1, 2, 3, \ldots, N$, where e is an error term. A set of N observations is simultaneously obtained for the X's and Y's; the remaining problem then is to estimate the b's. This equation asserts that a given value(s) of X can be multiplied by the estimated regression coefficient and added to an error term e to derive the corresponding Y value. The error term represents the discrepancy between the actual value of Y and that obtained from Xb. The better the model fit the smaller the e. The error term is incorporated in the model for three basic reasons. First, some factors may not be amenable to quantification or others may not be included because they have only a slight effect (Johnston 1972:10). Second, a basic and unpredictable element of randomness is present in the Y variable that can only be adequately handled by a random variable term. Third, there may be errors in the observation or measurement procedure.

Assumptions

The paramount assumption is that of linearity, which states that the regression equation should be linear in the unknown parameters. From a simple perspective this assumption can be checked by plotting one variable against another or by specifically testing for linearity. Direct examination of the residuals resulting from fitting a predictive model to a set of data can help to detect violations of this assumption. For example, the pattern of residuals may indicate that new terms should be added to an equation or that reexpression of the currently included terms would be helpful. If the relationship between a pair of variables is found to be nonlinear, it may be possible to make this relationship intrinsically linear through transformation.

Two other assumptions concern the independent variables. The first assumption is that the values of the explanatory or independent variables can be measured without error, which means that the act of sampling is the sole source of variation in the independent variables. The second assumption is that the values of the independent variables are fixed or nonstochastic. This is not the case for most independent variables used in predictive modeling, which are usually random. If the original assumption is replaced by the less restrictive assumption that the values are stochastic, most of the results of applying this technique (i.e., significance tests, confidence intervals, and so forth) will be valid, provided that the independent variables have a distribution function that does not involve the variance and that their distributions are independent of the error terms (Johnston 1972:267).

The assumptions incorporated into general linear regression also include the zero mean assumption, the constant variance assumption, and the independent error terms assumption. The zero mean assumption can be expressed as $E(e_i) = 0$, which means that the expected value of $\bar{e}_i$, the mean of the probability distribution of possible values of e_i, is zero. The constant variance or *homoscedastic* assumption is expressed $E(e_i^2) = \mathrm{Var}(e_i) = s^2$. When this assumption is not met, there is a constant form of error as a function of the independent variables. For example, the error may increase as the values of a variable in the equation increase. The independent error terms assumption, expressed $E(e_{ij}) = 0$ for $i \neq j$, simply means that autocorrelation is not a problem. This assumption need not be met simply to obtain estimates of b, but it is an important assumption when tests depending on assumed normality (e.g., t- or F-tests) are to be run or when confidence intervals are to be run based on the t or F distributions. An additional assumption about e_i is that the values of this variable will be normally distributed.

Finally, general linear regression techniques assume that the number of cases exceeds the number of variables and that multicollinearity (significant correlations among the predictor variables) is not a problem. Violation of either of these assumptions leads to a reduced rank dispersion or correlation matrix for the data (Cooley and Lohnes 1971:58–59; Tatsuoka 1971:130–135).

General linear regression, with all of its variations, is a very powerful statistical tool. Its strength in any given application for predictive modeling, however, depends on the assumptions that are fulfilled for that particular application. Some of the assumptions are more crucial than others, but it is desirable to know what consequences to expect when particular assumptions are not fulfilled, how to determine whether an assumption is satisfied, and what alternative methods to employ when the classical technique is inappropriate.

Violations of Some Assumptions

One of the most important assumptions of this technique is that no linear dependence exists among the explanatory variables. The least-squares estimator b requires the inversion of $X'X$, which is impossible if the rank of X, and hence the rank of $X'X$, is less than the number of variables. X is the data matrix, where each

row represents the observations on all of the variables for one case and each column represents the observations of one variable for all of the cases. Thus any element of the X matrix represents the value of one variable for one case. The $'$ (prime) indicates the transpose of the array that precedes it, i.e., the elements of the rows and columns are interchanged. Any beginning text on matrix algebra can be consulted for a more extensive explanation of these concepts. While this is a case of extreme multicollinearity that exists when some of the variables are perfectly correlated, a less extreme but still serious case arises when some of the variables are highly but not perfectly correlated.

Johnston (1972:160) has outlined some of the adverse consequences of collinearity, noting for example that estimation precision decreases so that it becomes difficult, if not impossible, to disentangle the influence of predictor variables. Individual estimates may be greatly in error and highly correlated with one another, and the variances of the coefficients will be large. A second adverse effect is that a variable may be incorrectly dropped from an analysis because its coefficient does not differ significantly from zero. It may be that the variable in question has no predictive power not because it is unrelated to the phenomenon being modeled but because it is highly correlated with another variable in the equation. Finally, under conditions of collinearity, estimates of coefficients become very sensitive to the particular data set, so that the addition of a few new observations produces large shifts in the coefficients.

One way around the multicollinearity problem is to do a principal components analysis on the set of independent variables (see previous discussion). When multicollinearity is a problem, a set of principal components that is smaller than the original set of variables will represent most of the variance. The scores on the eigenvectors can then be used as predictors. This represents a parsimony desirable in many scientific endeavors, whereby a reduced set of variables can represent the dominant patterns of covariation present in a data set. This solution also has a practical quality in addition to solving the collinearity problem—fewer degrees of freedom are used by the predictor variables.

Another crucial assumption of the linear regression model is that of zero covariance of the residuals. For a model with normally distributed residuals, this assumption implies that they are independent. In an ordinary least-squares context there are three main consequences of autocorrelated residuals. First, unbiased estimates of the coefficients may be obtained, but their error terms may be large compared with those achieved by alternative estimation techniques. Second, the variances of the coefficients may be underestimated, and the normal procedures for calculating the t- and F-tests may no longer be valid. Third, the predictions will be inefficient because their sampling variances will be large.

Regression models are probably some of the most widely used models in archaeological research. Most archaeologists are familiar with normal-theory regression models based on the general linear model and its attendant assumptions, including the one that requires continuous variables. Researchers may be less familiar with models in which the dependent response variable or one or more of the

explanatory variables are categorical, or with models in which the explanatory variables are a mixture of categorical and continuously distributed values. Such mixtures of categorical and continuous data often constitute the independent variables used in predictive models. It would be desirable, but also very unrealistic, for all independent variables used in a predictive model to be measured on an interval or ratio scale. More often than not, however, nominal and ordinal variables are theoretically as important as those measured at higher levels. Because nominal and ordinal variables are sometimes difficult to integrate with interval and ratio measurements, the former may be relegated to positions of lesser importance. Archaeologists commonly employ tests of association between nominal and ordinal variables; however, they rarely go one step further and use these distributions as the independent variable(s) to predict probabilities of group membership for a case. Multivariate logistic models offer a means of doing this.

Logistic Regression

The simplest categorized response variable is random, with only two possible outcomes—for example, the presence or absence of archaeological sites. Classic linear regression models will not work as a predictive mechanism using such variables. If we code the two possible outcomes of the categorized response variable as 1 and 0, representing the presence and absence, respectively, of a site in a sample unit, and then try to use such a response variable in a classic linear regression model with two explanatory variables X_{i1}, X_{i2}, in the general linear regression equation given above, we will face two major problems.

The first problem is a violation of the constant error variance assumption. This is because the error term e can have one of two possible values

$$e_i = 1 - (\hat{b}_0 + \hat{b}_1 X_{i1} + \hat{b}_2 X_{i2})$$

if $\hat{Y}_i = 1$, and

$$e_i = -(\hat{b}_0 + \hat{b}_1 X_{i1} + \hat{b}_2 X_{i2})$$

if $\hat{Y}_i = 0$.

Because the values of the response variable are binomially distributed, the two possible values of e_i occur with probabilities of p_i and $1-p_i$. The error variance is not constant but depends instead on the values of the independent explanatory variables. If ordinary least-squares estimation is used, the estimators of the b's are unbiased, but they are not the best minimum variance estimators of all linear unbiased estimators.

The second major problem concerns the predicted values that are generated if the normal regression model is used. $\hat{Y}_i$ can only have the values 1 and 0, and the expected value of $\hat{Y}_i$, $E(\hat{Y}_i)$, is a simple weighted average of the two possible values. The weights are given by the probabilities of the possible values. This is shown as

$$E(\hat{Y}_i) = [1 \times P_i] + [0 \times (1-P_i)] = P_i$$

where P_i is the probability that $\hat{Y}_i = 1$. From this equation,

$$E(\hat{Y}_i) = \hat{b}_0 + \hat{b}_1 X_{i1} + \hat{b}_2 X_{i2}$$

The problem is that the predicted values $\hat{Y}_i$ are interpreted as probabilities. They can take values between - infinity and + infinity and hence are unbounded, whereas probabilities are supposed to lie between 1 and 0. The predictions may therefore lie outside the range of probability and will be inconsistent with a probabilistic interpretation. From a modeling perspective, then, use of the normal linear regression model to analyze a categorized response variable causes problems. The linear logit model outlined below offers one possible solution to these problems.

The Logistic Model

A probabilistic interpretation of the regression model can be made only if P_i falls between 0 and 1. A simple model that can be used to provide a probabilistic interpretation is

$$p_i = \frac{e^{b_0 + b_1 X_{i1} + b_2 X_{i2}}}{1 + e^{b_0 + b_1 X_{i1} + b_2 X_{i2}}}$$

and

$$1 - p_i = 1/(1 + e^{b_0 + b_1 X_{i1} + b_2 X_{i2}})$$

These equations, which are nonlinear, can be rewritten as

$$\frac{p_i}{1 - p_i} = e^{b_0 + b_1 X_{i1} + b_2 X_{i2}}$$

A linear model can then be achieved by a logistic transformation p_i such that

$$\log_e \frac{p_i}{1 - p_i} = b_0 + b_1 X_{i1} + b_2 X_{i2}$$

If we let L_i equal the terms to the left of the equals sign we have a linear logistic model. Predictions from the linear logistic model can be written as

$$\hat{L}_i = \hat{b}_0 + \hat{b}_1 X_{i1} + \hat{b}_2 X_{i2}$$

While the predicted values can fall in the range from - infinity to + infinity, the predicted probabilities fall in the range from 0 to 1.

A least-squares estimation to solve for the parameters of the equation is accomplished by first replacing the probabilities to the left of the equals sign with observed relative frequencies, which are derived by grouping the observations into k sets. The model can then be written as

$$\bar{L} = log_e \frac{f_j}{1 - f_j} = b_0 + b_1 X_{j1} + b_2 X_{j2} + (\bar{L}_j - L_j)$$

Even with this model the error variances are still not constant across the j sets. Ordinary least-squares estimations of the b's can, however, be replaced by a set of weighted least-squares estimations. The weights are of the form $n_j f_j (1 - f_j)$, and they imply that as the number of cases in a set (n_j) increases, more weight is given to that set. Given n_j, as f_j approaches 0 or 1, less weight is given to these equations since $\bar{L}_j$ is very sensitive to small changes in mf_j and thus takes large negative or positive values. If f_j is either 0 or 1, $\bar{L}_j$ becomes infinitely large and cannot be accommodated; thus, it is excluded. Some researchers feel that this exclusion is a waste of information and advocate another form of weights so that the variables can be included. Solving a modified system of simultaneous equations (because of the weights), known as the normal equations, produces the required weighted least-squares estimations. Descriptions of the operations involved in solving the system of normal equations can be found in most textbooks on matrix algebra.

Testing the Fit of a Model

The fit of a particular model can be determined using a test statistic based on the weighted differences between the observed $\bar{L}_j$, based on relative frequencies, and $\hat{L}_j$, the predicted value. The test statistic

$$\sum_j (\bar{L}_j - \hat{L}_j)^2 W_j$$

follows a chi-square distribution with degrees of freedom equal to the number of sets minus the number of parameters estimated. If the test statistic is greater than the value shown on the table of chi-square values, the null hypothesis of no significant difference between the predicted and observed values is rejected, with the implication that the model does not represent the observed variation. If the test statistic yields a value that is lower than the critical chi-square value, the null hypothesis is accepted, with the implication that there is agreement between the predicted and observed logits.

In matrix notation the normal equation can be written as

$$(X' \cup^{-1} X)\hat{b} = X' \cup^{-1} \bar{L}$$

This equation can be solved for the b's as

$$\hat{b} = (X' \cup^{-1} X)^{-1} = X' \cup^{-1} \bar{L}$$

This also makes it easy to see how additional explanatory variables can be incorporated; all that is required is an additional column in the X matrix. The test statistic previously described is given in matrix terms as

$$(\bar{L} - X\hat{b})' \cup^{-1} (\bar{L} - X\hat{b})$$

Standard errors also provide a means to determine whether a parameter is significant. Standard errors are given by the square roots of the diagonal elements of the equation solved for the b's, above.

The validity of adding an additional explanatory variable c, or a set of variables, can be tested by considering the standard errors or the test statistic proposed by Grizzle et al. (1969), given as

$$\hat{b}'c'[c(X'\cup^{-1}X)c']^{-1}c\hat{b}$$

Under the null hypothesis, this test statistic has degrees of freedom equal to the number of rows of c. The null hypothesis tested by this statistic is that the additional parameter(s) is zero. If only one additional variable is added, c will be a vector of the form

$$c = [0_1 0_1 0_1 1]$$

The coefficient associated with c will be $cb = [0, 0, 0, 1]$ times a column vector with elements $[\hat{b}_0, \hat{b}_1, \hat{b}_2, \hat{b}_3]$.

To this point only a dichotomous response variable, such as the presence or absence of a site, has been considered. There are many more cases in which the archaeologist is faced with a categorized response variable with more than two possible outcomes. Such a variable is termed a *polychotomous* variable, and it probably more accurately represents situations that will be encountered in practice. An example of a variable with a multiple response category would be one with categories representing Pueblo I, Pueblo II, and Pueblo III sites. The linear logistic model can be extended to cover these cases, but its derivation is complex. We feel that this derivation is beyond the scope of this volume and refer the reader to Wrigley (1976) for a readable discussion of some of the math involved. If more than one explanatory variable is to be included in the extended model, all that is involved is the addition of an extra two columns in the matrix of observations, x. If the response variable has more than three outcomes an extra linear logistic equation is added to the pair of equations required for the three-outcome example. The other matrices—X, V^{-1}, I, and b—and vectors are also increased in size.

Maximum Likelihood Estimates

Wrigley (1976:27) notes that the problem with least-squares parameter estimation is that regrouping of the data set must be done for each model, which is quite a laborious procedure. The solution to this drawback is provided by maximum likelihood estimation. Determining the maximum of the log likelihood and the extension of this maximum to a multiple response category case requires a numerical optimization computer program. The logistic regression program PLR in BMDP is based on such a maximum likelihood estimation procedure. The B_i are estimated as the values that maximize the likelihood function using the method proposed by Jennich and Moore (1975).

Discriminant Function Analysis

Discriminant analysis is a broad term that refers to several closely related activities that can be divided into the processes of *(a)* defining the differences between groups and *(b)* classifying the cases into the groups. Groups are defined on the basis of a set of variables, and the variables with the most discriminating power are identified. For example, the groups could be different types of sites defined on the basis of a number of environmental and archaeological variables. The next step is to derive mathematical rules that can be used to classify the cases into the defined groups. Several different classification procedures exist, but all employ the concept of comparing a case's position with the centroids of the various groups to locate the closest centroid.

In the following section the general aspects of discriminant function analysis are outlined. Examples are drawn from a recent predictive modeling study of Fort Carson Military Reservation, Colorado (Altschul and Rose 1986). Fort Carson presented an interesting problem common to many CRM projects because it was not created with archaeology in mind and, therefore, its boundaries do not coincide with natural or cultural units. The area lies within the Arkansas River drainage. While several small drainages cut through the reservation, none are captured entirely within the base. After conducting several reservation-wide analyses we focused our attention on two large drainages, Turkey Creek and Red Creek. A predictive model developed for Turkey Creek successfully discriminated site locations from nonsites by classifying 82.4 percent of the sites correctly using a jackknife procedure (discussed below). All of the 11 sites classified as nonsites are located on or near Booth Mountain, a seemingly inhospitable uplift between Turkey Creek and Booth Gulch.

The Booth Mountain sites appear to be small, transient camps, perhaps indicating that this area was favored for certain resources. Such an explanation, while plausible, was difficult to accept given the present archaeological knowledge of the area. A separate predictive model was therefore developed for Booth Mountain in an attempt to distinguish between site and nonsite locations. The site group was composed of 17 locations, while the nonsite group consisted of 21 locations selected by a two-stage random sampling design. Because of the relatively small size of the analysis, Booth Mountain provides the opportunity to follow the discriminant procedure from beginning to end.

The means and standard deviations for variables used to characterize the sites and nonsites are presented in Table 5.3. As a group, sites are distinguished by being closer to the nearest water source, having a more southerly exposure, and commanding a wider view. The standard deviations of these variables are also smaller than those of their nonsite counterparts, indicating less variability about the average value—what would be expected if people were keying in on specific aspects of the environment when making locational decisions.

TABLE 5.3.

Means and standard deviations of variables for Booth Mountain nonsites, sites, and both groups combined

	Nonsites (n=21)	*Sites (n=17)*	*Combined (n=38)*
	Means		
VARIABLE			
RELIEF	99.04762	112.94118	105.26316
ASPECT	142.71429	189.47059	163.63158
DSTWTR	.17619	.11029	.14671
DSTRNK2	4.26381	4.65382	4.43829
ELVWAT	52.38095	61.76471	56.57895
VIEW	127.85714	187.17647	154.39474
	Standard Deviations		
VARIABLE			
RELIEF	39.23070	33.12188	36.64162
ASPECT	56.63934	48.45632	53.15818
DSTWTR	.14693	.07861	.12141
DSTRNK2	1.67176	1.83153	1.74458
ELVWAT	33.00072	47.10166	39.88804
VIEW	115.26677	69.65292	97.66056

Assumptions

The model on which the most common approaches to discriminant function analysis are based has a number of underlying assumptions. When the data do not satisfy the assumptions, the statistical results will not be an accurate reflection of the real world. First, the number of groups, the dependent variable, must be greater than or equal to two. In the Booth Mountain example we are attempting to discriminate between two groups, sites and nonsites. Second, there must be at least two cases per group. Realistically, there should be a large number of cases per group ($>$25) so that the sample statistics will accurately reflect the processes operating at the population level. Because the Booth Mountain model was primarily an exploratory attempt to see if patterns in site location could be discerned at a level that would warrant further work, we felt justified in relaxing this rule of thumb; for Booth Mountain there are 17 sites and 21 nonsites. Third, there should not be more discriminating variables than the total number of cases minus two. In the Booth Mountain example there are six variables: relief, aspect, distance to water, distance along the river to the nearest rank 2 stream, elevation above water, and view (see Altschul and Rose 1986 for operational definitions of these variables). Realistically, again, there should be many more cases than variables, say 10 times as many. Finally, the discriminating variables should be measured at the interval or ratio level of measurement. If categorical variables were used, especially those coded 0 and 1, multicollinearity would be a problem. All of the variables used in the Booth Mountain study are measured on at least an interval scale.

Three additional assumptions have to do with the logical and mathematical relationships among variables. The first of these is the assumption that multicollinearity, in the form of either linear combinations of variables or perfectly correlated variables, does not occur in the data. For example, if variables A through E were being considered for inclusion in an analysis, the sum of variables B and D, the product of variables A and B, or the average of variables B, C, D, and E are linear combinations that cannot be employed along with the original variables. This makes intuitive sense because variables that are linear combinations do not contain any information beyond what is offered by the individual variables. For similar reasons, when two variables are perfectly correlated, both cannot be used. If multicollinearity is a problem, some classification functions based on the concept of distance may be hard to define, and probabilistic interpretations associated with group membership may be difficult to formulate. If the variables are not significantly intercorrelated and they possess normal distributions, probabilities of group membership can be assigned. If significant multicollinearity exists, it may be necessary to drop a variable(s) from the analysis, create a hybrid variable that is based on several others, or use a data-reduction technique such as principal components analysis to express the original variables as uncorrelated linear combinations of one another.

Discriminant function analysis also assumes that the population covariance matrices for each group are approximately equal. The covariance between two variables is an unstandardized measure of how they vary *together*. Thus this measure can take on any range of values and is not, like the correlation coefficient, restricted to a particular range, e.g., between +1.0 and -1.0. The covariance matrix arises from the pair-wise arrangement of covariances into a table of rows and columns. The covariance matrix for the sample nonsite locations on Booth Mountain is shown in Table 5.4. The matrix has a row and a column for each variable, and the intersection of a row and a column contains the covariance for that pair of variables. Only the main diagonal and the lower left portion of the matrix are shown because the upper right portion is a mirror image of the lower left. For example, in Table 5.4 the covariance between aspect (row 2) and relief (column 1) is 213.7143, and the covariance between view (row 6) and elevation above water (column 5) is 687.8571. The covariance of a variable with itself is its variance; therefore, the diagonal from the upper-left corner to the lower-right corner contains the variances. The diagonal in Table 5.4 contains the variances of the six variables, with relief having a variance of 1539.048, aspect a variance of 3208.014, and so on for the remaining variables. If we have two or more groups, a covariance matrix can be computed from the cases in each group. The covariance matrix for the Booth Mountain sample of site locations is given in Table 5.5. Two or more covariance matrices are considered equal when the elements at similar positions in each matrix are not significantly different.

It is clear from an examination of Tables 5.4 and 5.5 that the group covariance matrices are unequal. Unequal group covariance matrices lead to canonical discriminant functions (discussed below) that may not yield maximum group separation and may result in unrealistic probabilities of group membership. Some of the error

TABLE 5.4.

Covariance matrix for Booth Mountain nonsite group

	RELIEF	*ASPECT*	*DSTWTR*	*DSTRNK2*	*ELVWAT*	*VIEW*
RELIEF	1539.048					
ASPECT	213.7143	3208.014				
DSTWTR	-1.248810	4.958357	.2158726E-01			
DSTRNK2	-4.606190	67.73639	.1034365	2.794780		
ELVWAT	-157.6190	711.7143	2.094524	9.950476	1089.048	
VIEW	92.85714	-771.3929	-.5785714	-20.61768	687.8571	13286.43

TABLE 5.5.

Covariance matrix for Booth Mountain site group

	RELIEF	*ASPECT*	*DSTWTR*	*DSTRNK2*	*ELVWAT*	*VIEW*
RELIEF	1097.059					
ASPECT	-535.2206	2348.015				
DSTWTR	.4897059	-.8567096	.6179596E-02			
DSTRNK2	-32.25882	62.16465	.2590993E-02	3.354502		
ELVWAT	263.2353	-1030.257	3.202574	-16.55092	2218.566	
VIEW	261.3235	-612.7757	1.460570	-27.32134	1138.732	4851.529

results from the calculation of the within-groups covariance matrix, which is supposed to be an estimate of the common equal group covariance matrices in the population. One way of coping with the problem of unequal covariance matrices is quadratic discrimination, which bases the probability of group membership on the individual group covariance matrices. Quadratic procedures require larger sample sizes because more terms are added to the equation, and thus could be difficult to use in predictive modeling situations, where samples are frequently small. In the case of the Booth Mountain data (with group sizes of 17 and 21) quadratic discrimination would be inappropriate (and therefore is not discussed further here). In this case a linear discriminant function like the canonical discriminant function discussed below is a better approach, although we still have to accept some degree of error in the within-groups covariance matrix. Fortunately, this assumption has received considerable theoretical and applied scrutiny, most of which points to the fact that this assumption can be relaxed.

A final assumption is that each group represents a population with a multivariate normal distribution on the discriminating variables. A multivariate normal distribution exists when each variable has a normal distribution when the values of the other variables are held constant. This assumption makes it possible to compute meaningful tests of significance and probabilities of group membership. This

normality assumption is required by most of the popular computer packages, such as SPSS (Nie et al. 1975) and BMDP. However, it should be noted that discriminant analysis can be performed using other parametric distributions (though not with SPSS or BMDP), and that nonparametric discriminant techniques that do not assume a normal distribution could possibly be employed. Nonparametric discriminant analysis computer programs and discriminant analysis procedures based on other parametric distributions are not as widely available as their normal theory counterparts.

Canonical Discriminant Functions

A discriminant function is a linear combination of discriminating variables formed to satisfy certain conditions (Klecka 1980:15). The form of the discriminant function in summation notation is

$$f_{km} = u_0 + u_1X_{1km} + u_2X_{2km} + \ldots + u_pX_{pkm}$$

where f_{km} is the score on the discriminant function for case m in the kth group, X_{km} is the value on the discriminating variable X_i for the mth case in the kth group, and u_i is the coefficient of the function derived according to certain characteristics.

The maximum number of unique functions that can be derived in an analysis is equal to either one less than the number of groups or to the number of discriminating variables, whichever is less. If there are three groups, we can derive two discriminant functions; if there are only two groups, as in the Booth Mountain example, then only one discriminant function can be derived.

Coefficients (u's) for the first function are derived in such a way that the group means on the function are as different as possible—in other words, so that group differences are maximized. Coefficients for the second function are derived such that the differences among group means are maximized under the additional constraint that these coefficients are uncorrelated with those of the first function. Additional functions continue to be derived so that group differences are maximized and the coefficients are uncorrelated with those of the previous functions.

Many multivariate statistics texts discuss the mathematical aspects of discriminant function analysis; particularly useful and readable examples are Cooley and Lohnes (1971), Harris (1975), Morrison (1976), and Tatsuoka (1971). We will simply review in the most general fashion some fundamental principles underlying the derivation of canonical discriminant functions. The first requirement is to assess the degree of differences among the data cases. This is done with the sums-of-squares-cross-products matrix (SSCP), where the value of a particular element (t_{ij}) is given as

$$t_{ij} = \sum_{k=1}^{g} \sum_{m=1}^{n_k} (X_{ikm} - X_{i\cdot\cdot})(X_{jkm} - X_{j\cdot\cdot})$$

where g is the number of groups, n_k is the cases in group k, $n.$ is the total number of cases in all groups, X_{ikm} is the value of variable i for the mth case in the kth group, X_{ik}.

is the mean for variable i for all the cases in the kth group, and $X_{i..}$ is the mean for variable i for all the cases in all the groups.

The first set of values in parentheses is the amount by which the value of a case deviates from the grand mean of variable i. The second set of values in parentheses is the same information but for variable j. Each element of the diagonal in the SSCP matrix is simply the sum of squared deviations from the grand mean, since when i equals j the two terms are the same. If i does not equal j the result is the sum of a deviation on one variable multiplied by the deviation on the other. This is a way to measure the covariation (correlation) between two variables, since it tells us how the magnitude and direction of a deviation on one variable correspond to those on another. The covariance matrix, the subject of much discussion in discriminant function analysis, is produced from the SSCP matrix by dividing each element by (n.-1). The covariance matrix for the Booth Mountain analysis is shown in Table 5.6. This particular covariance matrix is called a total covariance matrix, since it is based on the cases from both the site and the nonsite groups. A covariance matrix can be calculated for each group if only the cases assigned to that group are used (see the individual group covariance matrices presented in Tables 5.4 and 5.5).

TABLE 5.6.

Total covariance matrix based on all Booth Mountain cases, both nonsites and sites

	RELIEF	*ASPECT*	*DSTWTR*	*DSTRNK2*	*ELVWAT*	*VIEW*
RELIEF	1355.334					
ASPECT	49.01849	3304.509				
DSTWTR	-.6957326	1.527404	.1544362E-01			
DSTRNK2	-15.06373	68.12651	.5050638E-01	2.999910		
ELVWAT	61.73542	50.59744	2.360064	-.8492532	1570.413	
VIEW	372.4609	22.28450	-.6736664	-17.08498	1005.576	10173.27

Another simple operation on the SSCP matrix converts it to a correlation matrix. Because the correlation coefficient is standardized to vary between +1.0 and -1.0, it is easier to understand as a measure of association between two variables than the covariance matrix. The elements of the SSCP matrix are converted to correlations by dividing each element by the square root of the product of the two diagonal elements falling in the same row and column. A similar operation can be used to convert the covariance matrix to a correlation matrix.

When the groups under consideration are distinct, the variability within the groups will be less than that between the groups. The degree of variability within each group is measured by a matrix called the within-group sums-of-squares-cross-products matrix (W), which is very similar to the SSCP. Unlike the SSCP, the deviations are measured from the mean of the group to which a case belongs. Elements of W are defined as

$$w_{ij} = \sum_{k=1}^{g} \sum_{m=1}^{n_k} (X_{ikm} - X_{ik\cdot})(X_{jkm} - X_{jk\cdot})$$

As with the SSCP matrix, the elements of W can be converted to a within-group covariance matrix by dividing each by ($n.$-g). The within-groups covariance matrix for the Booth Mountain analysis is given in Table 5.7. The W matrix can also be converted to a within-groups correlation matrix. The within groups correlation matrix for Booth Mountain is presented in Table 5.8. Each correlation measures the pair-wise correlation of variables within the groups and will usually differ from the total correlation, which is affected by group differences. If the data cases come from group populations with similar covariance structures, the within-groups correlations will provide a better estimate of the relationships between variables than the total correlations.

TABLE 5.7.

Booth Mountain within-group covariance matrix

		RELIEF 2	*ASPECT* 3	*DSTWTR* 4	*DSTRNK2* 5	*ELVWAT* 6	*VIEW* 8
RELIEF	2	1342.60815					
ASPECT	3	-119.14566	2825.79225				
DSTWTR	4	-.47614	2.37388	.01474			
DSTRNK2	5	-16.89625	65.26006	.05862	3.04355		
ELVWAT	6	29.42733	-62.49533	2.58699	-1.82792	1591.05587	
VIEW	8	167.73109	-700.89636	.32771	-23.59708	888.24580	9537.58450

TABLE 5.8.

Booth Mountain within-group correlation matrix

		RELIEF 2	*ASPECT* 3	*DSTWTR* 4	*DSTRNK2* 5	*ELVWAT* 6	*VIEW* 8
RELIEF	2	1.00000					
ASPECT	3	-.06117	1.00000				
DSTWTR	4	-.10703	.36783	1.00000			
DSTRNK2	5	-.26432	.70370	.27675	1.00000		
ELVWAT	6	.02013	-.02947	.53421	-.02627	1.00000	
VIEW	8	.04687	-.13501	.02764	-.13850	.22802	1.00000

If the centers of the groups are in the same location, the elements of the SSCP matrix will equal the elements of the W matrix. If the centers of the groups are different, the elements of the SSCP matrix will be larger than the elements of W. The difference between the SSCP and W matrices is measured by the between-groups sums-of-squares-cross-products matrix (B). The W and B matrices contain all the information about the relationships within the groups and between them.

The following set of simultaneous equations is solved for the values of b_{pi} and v.

$$\Sigma\, b_{1i}v_i = \lambda \Sigma\, w_{1i}v_i$$
$$\Sigma\, b_{2i}v_i = \lambda \Sigma\, w_{2i}v_i$$
$$\vdots \qquad \vdots$$
$$\Sigma\, b_{pi}v_i = \lambda \Sigma\, w_{pi}v_i$$

Lambda (λ) is called the eigenvalue, the v's are a set of p coefficients, and the b's and w's are previously defined quantities calculated from the sample data. The equations are solved subject to the constraint that the sum of the squared values of the v's must equal 1. Each unique nontrivial solution, with its lambda and set of v's, corresponds to one canonical discriminant function. The v's cannot be interpreted because the solution to the equations places no restriction on the origin or measurement units used for the discriminant space. Also, the scores produced for each case have no meaning. The discriminant space yields maximum separation between groups, but the groups can be anywhere in the space.

The u's of the first equation are given as follows:

$$u_i = v_i\sqrt{n. - g}\ , \quad u_0 = -\sum_{i=1}^{p} u_i x_{i..}$$

Using the u's gives discriminant scores (f's) for the cases that are in standard form, but the coefficients are regarded as unstandardized because the original data are not standardized. The unstandardized discriminant function coefficients from the first (and only) discriminant function calculated for the Booth Mountain analysis are given in Table 5.9. To calculate a score for any case for this discriminant function the actual data values for aspect, distance to water, and elevation above water would be multiplied by their respective coefficients and summed, along with the constant value. The scores for all the cases will then have a mean of 0 and a standard deviation of 1.0. The score for a case shows where it is on the axis defined by the function. Employing the u's instead of the v's does not change the amount of discrimination nor the relationship among groups. It does, however, move the origin of the discriminant function axes to coincide with the grand centroid, the point where all

TABLE 5.9.

Booth Mountain unstandardized discriminant function coefficients

	Function 1
ASPECT	.1819108E -01
DSTWTR	-9.021173
ELVWAT	.1901548E -01
(CONSTANT)	-2.729011

the discriminating variables have their average values over all cases. This relocation makes it possible to see how a group centroid or an individual case is located relative to the center of the system. The adjusted coefficients also produce discriminant scores measured in standard deviation units. Thus, if a case had a score of +2.9 we would know that it was distant from the center.

While the unstandardized coefficients tell us the absolute contribution of a variable in figuring the discriminant score, standardized coefficients must be used if we wish to determine the relative importance of a variable. The above equation would have yielded standardized coefficients if the original data had been in z-score form. Standardized coefficients (c's) can also be computed from the unstandardized coefficients (u's) by

$$c_i = u_i \sqrt{\frac{w_{ii}}{n. - g}}$$

where w_{ii} is the sum of squares for the variable i, n is the total number of cases, and g is the number of groups. The standardized discriminant function coefficients for the Booth Mountain example are given in Table 5.10.

The larger the magnitude of the standardized coefficient, disregarding the sign, the greater a particular variable's contribution to the discriminant score. From Table 5.10 we can see that aspect and distance to water are weighted about the same and are more important than elevation above water in discriminating between sites and nonsites. Scores from standardized coefficients can be computed by multiplying them by the data in z-score form, but scores are usually computed from the raw data values and the unstandardized coefficients, while the standardized coefficients are used to assess the relative importance of a variable. The limitation on this standardized coefficient is that if two variables are highly correlated they will share the same discriminating information and, hence, will share the contribution to the calculation of the score. For this reason the standardized coefficients for the two correlated variables may be smaller than they would be if only one of the variables was used. The individual variable coefficients might also be large but have opposite signs, so that the contribution of one is cancelled by the contribution of the other. Structure coefficients, discussed later, are not affected by relationships with other variables.

TABLE 5.10.

Booth Mountain standardized discriminant function coefficients

	Function 1
ASPECT	.96700
DSTWTR	-1.09522
ELVWAT	.75849

After the discriminant functions are computed they can be interpreted by considering *(a)* the relative positions of data cases and group centroids and *(b)* the relationships between the individual variables and discriminant functions. As previously described, the discriminant scores are computed by taking the original value for a case on each variable and multiplying it by the coefficient for that variable—then the products are added along with the constant term. The constant adjusts for the means, so the mean discriminant score will be zero over all cases. The unstandardized coefficients represent the amount of change in a case's position on that function if its score on the corresponding variable changed by one unit. Case-by-case inspection of a large number of cases usually is not that informative, though it may help in delineating outliers. Instead, it is more informative to focus on the group centroids, which are calculated by using the group means in the formulas. Sometimes it is easier to visualize the group centroids by plotting the data cases.

When there are two groups, such as sites and nonsites, there is only one discriminant function for which data can be plotted. In this situation the data cases can be arranged along a straight line to show what part of the function is "occupied." A better strategy is to construct a histogram for each group, with the continuum divided into intervals of standard deviation units, such as 0.10, 0.15, 0.20, 0.25, . . ., or whatever seems reasonable. Visual inspection of the histograms allows us to assess the density and distribution of each group and the relative group locations. The histograms for the nonsite and site groups are shown in Figures 5.9 and 5.10, respectively. The "stacked" histogram for both groups is shown in Figure 5.11. These histograms indicate that nonsite locations are more scattered than locations in the site group, with respect to the canonical discriminant function. The center of the nonsite group is at -0.726, while the center of the site group is at 0.897.

The locations of the group centroids and data cases can be plotted in an x,y coordinate system when there are two discriminant functions. A three-dimensional plot could also be prepared for a three-function situation, but four or more cannot be represented. In the latter situation, however, the first two functions are the most powerful discriminators, and a plot based on these two alone could be very informative. Two-dimensional plots are helpful when there is little overlap among the groups. If the groups are less distinct, and especially if there are a large number of cases, the plot may be difficult to interpret. In such a situation a plot of only the centroids or separate plots for each group may be more helpful.

The similarity between the discriminant function and a single variable can be assessed by computing the product-moment correlation between the two. These correlations are called *structure coefficients*. When the value of the coefficient is near zero, the variable and the function have little in common. When the value of the coefficient is very large, near +1.0 or -1.0, the function and the variable are highly correlated. The coefficients make it possible to "name" a discriminant function by noting the variables with which it is most highly correlated. The function can then be named after the characteristic(s) defined by the particular variable(s). For example, in a predictive modeling situation the first discriminant function might represent a configuration of variables depicting the relationship between site location and distance to water.

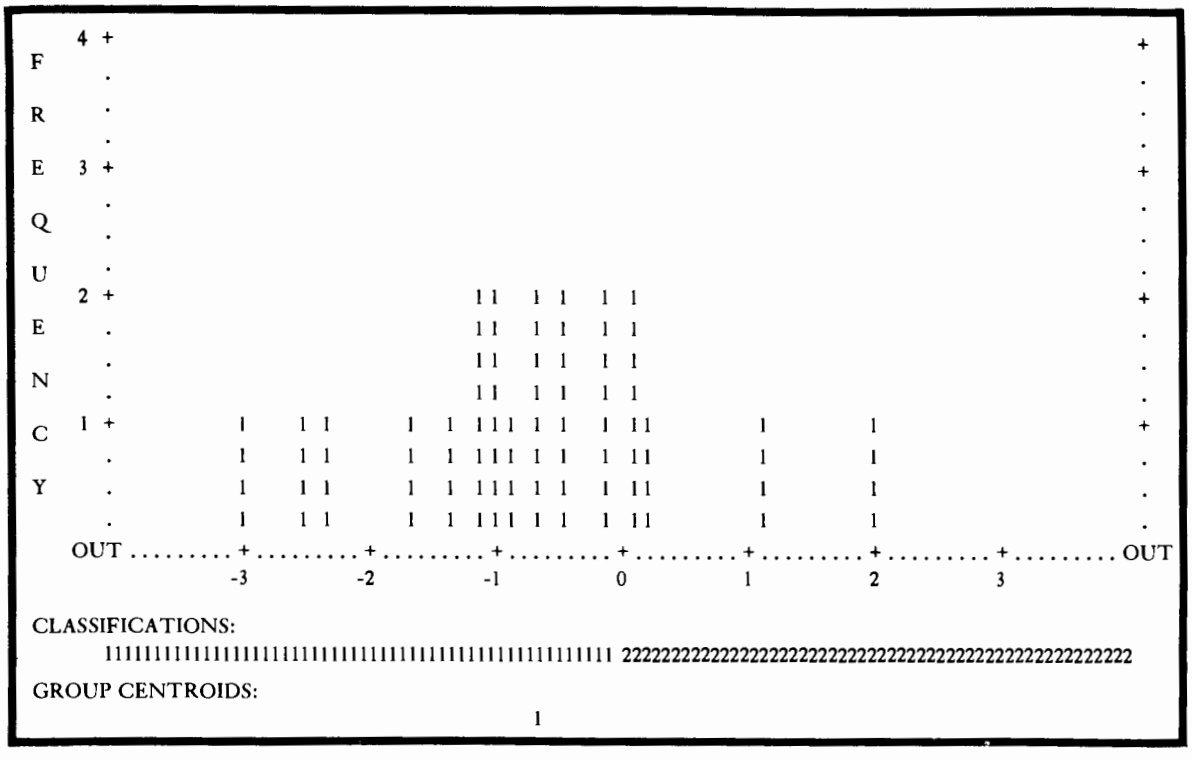


Figure 5.9. Histogram of discriminant function scores for Booth Mountain nonsite locations.

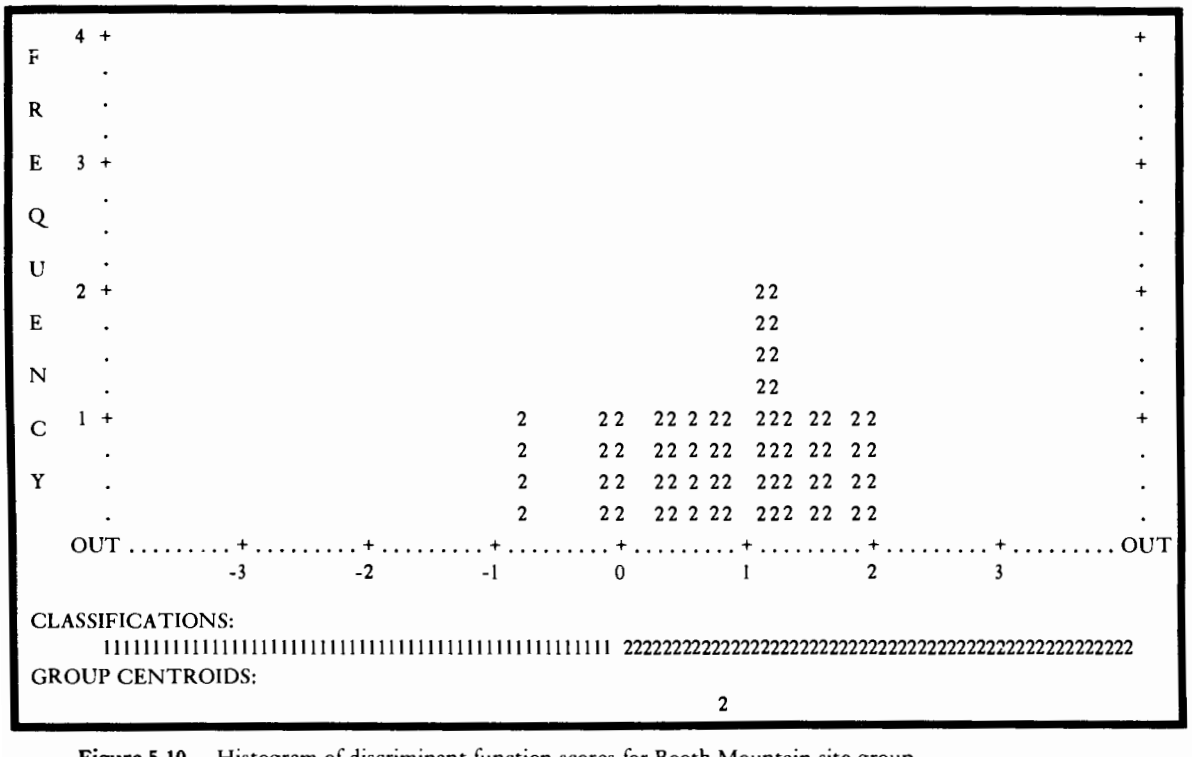


Figure 5.10. Histogram of discriminant function scores for Booth Mountain site group.

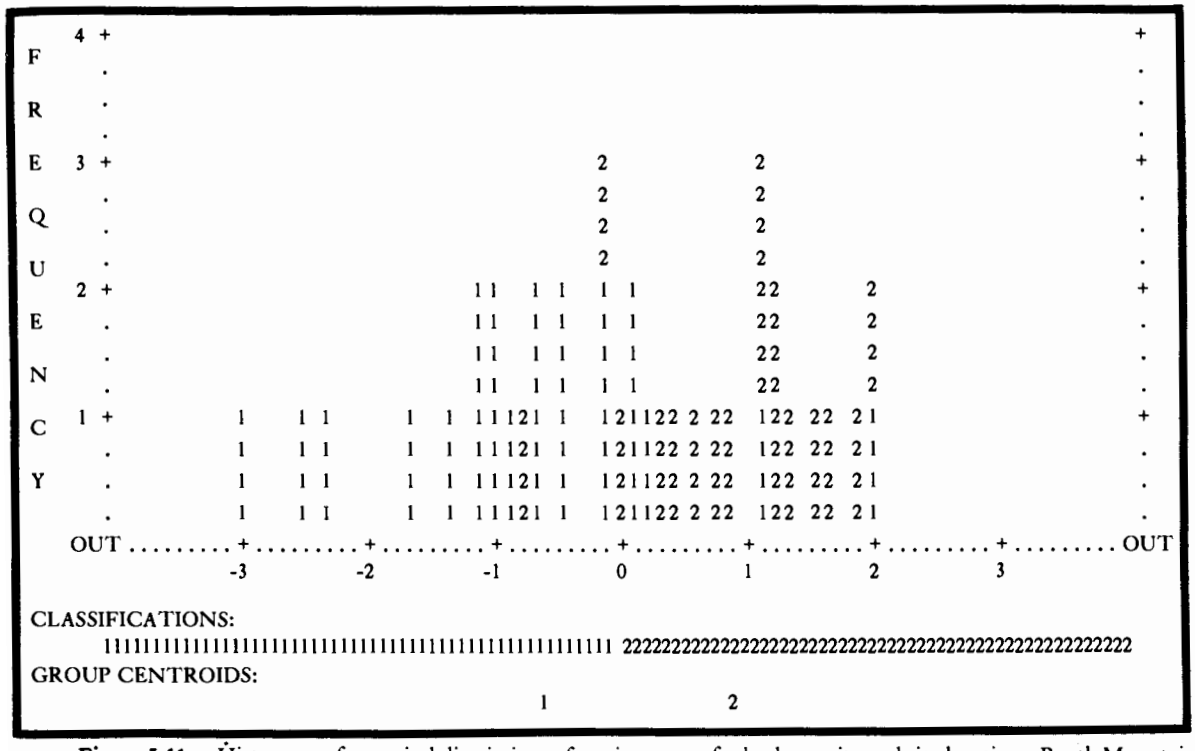


Figure 5.11. Histogram of canonical discriminant function scores for both nonsite and site locations, Booth Mountain.

These structure coefficients are more accurately called *total structure coefficients* since they portray the information carried by the discriminant functions that is useful in discriminating between groups. If we want to know how the functions are related to variables within the groups, however, pooled within-groups correlations are required. Known as *within-groups structure coefficients*, these are calculated as

$$s'_{ij} = \sum_{k=1}^{p} r'_{ik} c_{kj} = \sum_{k=1}^{p} \frac{w_{ik} c_{kj}}{\sqrt{w_{ii} w_{kk}}}$$

where s'_{ij} is the within-groups structure coefficient for variable i and function j, r'_{ik} *is the pooled within-groups correlation between variables* i and k, and c_{kj} is the standardized discriminant function coefficient for variable k on function j. The pooled within-groups correlations between the Booth Mountain canonical discriminant function and the variables are given in Table 5.11.

It is possible for a variable to have a low standardized coefficient in a discriminant function but a relatively large total structure coefficient because the structure coefficients are bivariate correlations that are not affected by relationships with

TABLE 5.11.

Pooled within-group correlations between the canonical discriminant function and the discriminating variables

	Function 1
ASPECT	.54179
DSTRNK2	.35745
DSTWTR	-.33434
ELVWAT	.14491
RELIEF	.07335
VIEW	.01212

other variables. It is also possible for two variables to have large standardized coefficients of opposite signs but small structure coefficients; when the discriminant function scores are calculated these variables tend to cancel each other out. For these reasons structure coefficients are a better indicator of the meaning of a discriminant function than are the standardized coefficients.

When the number of groups and variables is large there will be a number of discriminant functions, not all of which will be nontrivial or statistically significant (Klecka 1980:34). In an effort to determine how many discriminant functions to retain we can look at the solutions of the functions. When each discriminant function is solved, an eigenvalue (λ) and a set of coefficients are produced. The eigenvalues will be positive or zero; the larger the lambda, the more the groups will be separated on that function. The eigenvalue associated with the discriminant function derived in the Booth Mountain example is 0.69. We can also determine the total discriminating power for each discriminant function by converting the eigenvalues to relative percentages. Each eigenvalue is divided by the sum of all the eigenvalues and multiplied by 100. These relative magnitudes make it easier to see the discriminating power of each function. In our example, only one discriminant function can be derived so it represents 100 percent of the discriminating power.

A third way to judge the utility of a discriminant function is by examining the canonical correlation between the groups and the function. The canonical correlation coefficient describes the relationship between two separate sets of interval-level variables. It ranges from 0 to +1, with zero indicating no association between the groups and the discriminant function and larger numbers representing increasing degrees of association. The canonical correlation in our Booth Mountain example is 0.64. The canonical correlation (r^*) is related to the eigenvalue by

$$r^* = \sqrt{\frac{\lambda_i}{1 + \lambda_i}}$$

where i denotes the ith discriminant function.

Another means of evaluating the utility of the discriminant functions is to test the statistical significance of discriminating information not already accounted for

by the earlier functions. This can be done using Wilks's lambda, a multivariate measure of group differences over several variables. Starting with no functions, Wilks's lambda is calculated as each is derived. In the Booth Mountain analysis, Wilks's lambda associated with the first discriminant function is 0.592. Lambda is an inverse measure of the discriminating power of the variables that have not been removed by the discriminating functions; the larger the value, the smaller the amount of information remaining.

Wilks's lambda is actually more useful as an intermediate statistic than as an end product because the results can also be converted into a test of significance. This is easily accomplished by converting lambda to an approximation of either the chi-square or F distribution. Chi-square, the easier of the two to calculate, is based on the following formula:

$$\chi^2 = -[n. - \left(\frac{p+g}{2}\right) - 1] \; log_e \; \Lambda \; k$$

with $(p\text{-}k)(g\text{-}k\text{-}1)$ degrees of freedom. The chi-square value can be compared with standard tables to determine the significance level. Many computer programs print the exact significance level. The chi-square value associated with the Booth Mountain Wilks's lambda of 0.592 is 18.058, with 3 degrees of freedom and a probability of 0.0004. Thus, the group differences will be significant before any discriminant functions have been derived (that is, when $k = 0$). After the first function has been derived, a check is made to see if any remaining discrimination is significant. If it is, a second function is derived and the test is made again. This process is continued until Wilks's lambda is not significant. Of course, in this example only one function can be derived.

The point must be made that assessments of statistical significance are appropriate only when the data are derived from a sample with a probabilistic basis (see Chapter 6). Generally, simple random sampling is assumed since the derivation of tests for other sampling situations is more complex. When the cases have been procured by some procedure other than simple random sampling, tests should be interpreted conservatively.

Classification

Earlier the purpose of discriminant function analysis was divided into two parts, derivation of the discriminant functions and classification. By classification we mean the process of identifying the likely group membership of a case when the only information known is the case's values on the discriminating variables. Classification procedures can use either the discriminating variables or the canonical discriminant functions. Klecka (1980:42) notes that in the first instance a discriminant analysis is not even performed. Classification functions are derived using the theory of maximum group differences and tests are not made for the significance of the discrimination or the dimensionality of the discriminant space. When canonical

discriminant functions are derived first and classification is based on these functions, a more thorough analysis can be performed.

Fisher (1936) first suggested that classification should be based on a linear combination of the discriminating variables. His linear combination was based on maximizing differences between groups and minimizing variation within groups. Traditional classification functions are based on the pooled within-groups covariance matrix and the centroids of the discriminating variables. They have the form

$$b_k = b_{k0} + b_{k1}X_1 + b_{k2}X_2 + \ldots + b_{kp}X_p$$

where b_k is the score for group k and the b's are coefficients that need to be derived. There is always a separate equation for each group; if there are four groups, each case will have four scores. Fisher's linear discriminant functions for the Booth Mountain nonsite and site groups are given in Table 5.12. A particular case is classified into the group with the highest score. A straightforward application of this classification procedure results in the correct group placement (nonsites/sites) of 86.8 percent of the 38 locations. Of the nonsites, 90.5 percent (19 of 21) are correctly classified, while 82.4 percent (14 of 17) of the sites are correctly classified.

Another means of classification is to measure the distances from the individual case to each group centroid and to assign the case to the closest group. Mahalanobis (1963) proposes a generalized distance measure that circumvents the problems of intercorrelations among variables that do not have the same units of measurement or standard deviations. The generalized distance measure is given as

$$D^2(X|G_k) = (n. - g) \sum_{i=1}^{p} \sum_{j=1}^{p} a_{ij}(X_i - X_{ik\cdot})(X_j - X_{jk\cdot})$$

where $D^2(X|G_k)$ is the squared distance from case X to the centroid of group k. The case is then classified into the group with the smallest D^2 value. The formula given above assumes equal group covariance matrices, but Tatsuoka (1971:222) gives a modified form of the equation for unequal group covariance matrices. D^2 has the properties of the chi-square statistic with p degrees of freedom. Thus, the distance is measured in chi-square units, and if each group comes from a population with a multivariate normal distribution, most of the cases will be clustered near the centroid. The density of cases decreases farther away from the centroid.

TABLE 5.12.

Classification functions for the Booth Mountain nonsite and site groups (Fisher's linear discriminant functions)

	Nonsites	*Sites*
VARIABLE		
3 ASPECT	.05504	.08457
4 DSTWTR	-4.29467	-18.94006
6 ELVWAT	.04207	.07294
CONSTANT	-5.26727	-9.99673

These distances can be converted into probabilities of group membership if the assumption of multivariate normality is met (Nie et al. 1975:445). Assigning a case to the group with the highest score is then equivalent to assigning a case to the group for which the probability of group membership is highest. Thus far it has been assumed that each group is treated equally. A Bayesian adjustment of the probability of group membership, however, is often desirable (see below).

The canonical discrimination functions can be used for classification in place of the original discriminating variables. Klecka (1980:47) notes that when there are a large number of cases to be classified, the canonical functions make the task easier; in the first case distances for p variables must be computed, while the second case requires only computation of q discriminant functions. The classifications based on the canonical functions will usually be the same as those using the original discriminating variables. One exception to this is when the group covariance matrices are not equal because the canonical discriminant functions use the within-groups SSCP matrix, which is the weighted average of the individual SSCP matrices. Tatsuoka (1971:232–233) notes that the two classification procedures yield closely similar results and that the canonical procedure can be used unless the differences among the group covariance matrices are drastic. The two procedures may also give different results when one or more canonical discriminant functions are dropped because they are not statistically significant. In this instance, the canonical results should be more accurate than classification based on the original variables because the effect of idiosyncratic sample variability is reduced.

Classification of the cases in the Booth Mountain analysis is presented in Table 5.13. In addition to the actual group membership and discriminant scores computed from the canonical discriminant function, several probability values are presented for the highest and second-highest probability groups. $P(X|G)$ is the probability that a case that far from the center of the indicated group (0 = nonsites, 1 = sites) would actually belong to it, while $P(G|X)$ is the probability that the particular case is a member of the indicated group. For example, the first case in Table 5.13, a nonsite, is correctly placed in group 0 (nonsites). The probability that a location that far from the nonsite group centroid actually belongs in the nonsite category is 0.5098, while the probability that the location is a member of the nonsite group is 0.9275. The probability that the location is a member of the second highest probability group, the site group, which in this case is the *only* other group, is 0.0725. The probabilities of group membership of the other cases can be considered in a similar manner. Incorrect group classifications of cases are denoted with asterisks (***) following the actual group listing.

To this point all groups have been considered as if they were of equal size. In some situations this may be unrealistic, as when 90 percent of the locations belong in group A and only 10 percent belong in group B. We do not need to do a discriminant analysis to know that in this situation there is a high probability that any given case will belong to group A, and that we would only want to place a case in group B if there was a strong reason for doing so. In the Booth Mountain analysis there are a total of 38 locations, 21 nonsites and 17 sites. Thus, if a location was randomly

TABLE 5.13.

Booth Mountain discriminant scores and probabilities of group membership. P(X | G) is the probability that a case that far from the group centroid belongs to the group, while P(G | X) is the probability that the case belongs to the indicated group.

Case Sequence Number	*Actual Group*	*Highest Probability*			*Second Highest Probability*		*Discriminant Scores*
		Group	*P(X | G)*	*P(G | X)*	*Group*	*P(G | X)*	
1	0	0	.5098	.9275	1	.0725	-1.3855
2	0 ***	1	.2717	.9499	0	.0501	1.9963
3	0	0	.5296	.6125	1	.3875	-.0977
4	0	0	.8591	.8540	1	.1460	-.9039
5	0	0	.8342	.7574	1	.2426	-.5169
6	0	0	.3674	.5037	1	.4963	.1752
7	0	0	.0215	.9946	1	.0054	-3.0261
8	0	0	.0636	.9889	1	.0111	-2.5811
9	0	0	.4097	.5349	1	.4651	.0982
10	0	0	.8054	.8674	1	.1326	-.9726
11	0	0	.8198	.7518	1	.2482	-.4984
12	0	0	.5614	.6307	1	.3693	-.1456
13	0	0	.7836	.8726	1	.1274	-1.0009
14	0	0	.0933	.9853	1	.0147	-2.4046
15	0	0	.3933	.5231	1	.4769	.1273
16	0 ***	1	.8188	.8219	0	.1781	1.1262
17	0	0	.9822	.8088	1	.1912	-.7040
18	0	0	.6979	.8917	1	.1083	-1.1145
19	0	0	.9874	.8104	1	.1896	-.7105
20	0	0	.7443	.8816	1	.1184	-1.0524
21	0	0	.3502	.9523	1	.0477	-1.6604
22	1	1	.2537	.9531	0	.0469	2.0387
23	1	1	.8256	.8198	0	.1802	1.1175
24	1	1	.9776	.7691	0	.2309	.9253
25	1	1	.7330	.8470	0	.1530	1.2383
26	1 ***	0	.9422	.8314	1	.1686	-.7987
27	1	1	.8001	.6784	0	.3216	.6439
29	1	1	.6893	.8589	0	.1411	1.2970
30	1	1	.3372	.9379	0	.0621	1.8569
31	1	1	.8378	.8161	0	.1839	1.1019
32	1	1	.4595	.9136	0	.0864	1.6368
33	1	1	.6100	.5816	0	.4184	.3871
34	1 ***	0	.5140	.6032	1	.3968	-.0737
35	1	1	.5822	.5657	0	.4343	.3471
36	1	1	.7523	.8415	0	.1585	1.2128
37	1	1	.9089	.7254	0	.2746	.7827
38	1	1	.5398	.8959	0	.1041	1.5104
39	1 ***	0	.4506	.5630	1	.4370	.0281

*** incorrect classification

Note: Case 28 is a historical site that was not included in the analysis.

selected it would have a slightly higher chance (0.54 vs 0.46) of being placed in the nonsite group simply because there are more nonsites than sites. This problem can be solved by adjusting the posterior probability of group membership to account for the prior knowledge that we have about group membership. Prior knowledge can be regarded as any information about the populations that does not result from the current research but that permits us to formulate hypothesized probabilities of group membership for a randomly selected case (Cooley and Lohnes 1971:263). Classification decisions involving a priori probabilities of group membership are examples of Bayesian logic. Adjustment of a classification function to take prior probabilities into account may improve prediction accuracy and help to minimize the cost of making errors. If groups are well separated, the use of prior probabilities is unlikely to affect the results of the classification; use of these probabilities is most effective when group separation is weak.

Classification boundary lines can be superimposed on one- and two-dimensional plots of the cases. These lines partition the measurement space into group territories within which individual cases are classified. When discrimination is weak, many cases may fall outside their group's territory and will be misclassified. If there is only one function, the dividing point between the two groups is one-half the sum of the discriminant scores for the two group centroids. With two functions we solve for $D^2(X|G_i) = D^2(X|G_k)$, which results in the equation for a straight line when the group covariance matrices are equal. With unequal covariances the boundary in the two-function case will be a curve around the group with the smaller dispersion (Van de Geer 1971:263–265). In the single-function case the dividing point will be closer to the group with the smaller variance.

A final consideration in discriminant analysis involves selecting the variables to be included in the discriminant function. It is possible to use the entire set of independent variables, regardless of the discriminating power of each. This approach may be appropriate for theoretical reasons if priorities cannot be placed on specific variables. In other situations, theory may not provide a strong enough reason for specifying the exact list of discriminating variables. Theoretical knowledge may merely suggest potential discriminating variables, or the investigation may be exploratory and the discovery of discriminating variables may be a prime objective of the research. In these situations some variables may not produce good group separation because the group means are too similar. Alternatively, two or more variables may possess similar discriminating information and hence be redundant.

A stepwise selection procedure can be used to select the most powerful discriminating variables. There are two means of doing this: *forward selection* begins by selecting the single variable that produces the greatest univariate discrimination. The first variable is then paired with each of the remaining variables, one at a time, to find the pair offering the greatest group discrimination. The first and second variables are then combined with each variable in turn to form a triplet. This process continues until all possible variables have been selected or until those remaining do not contribute significant discrimination. Output from a stepwise

selection procedure, used in the Booth Mountain example, is shown in Table 5.14. It will be considered in more detail below.

In the *backward selection* procedure all variables are used at the beginning, and the one with the weakest discriminating power is eliminated at each step. Forward and backward selection procedures can also be combined. Generally this involves a forward selection procedure, in which each step begins with a review of the remaining variables in the equation. When a variable no longer makes a significant contribution, it is deleted; it may, however, be selected again at a later step. The order in which variables are selected does not always coincide with their relative importance. Because of intercorrelations, certain variables may be prevented from entering the calculations.

Before a variable is tested on a selection criterion many computer programs require it to pass certain minimal conditions to ensure accuracy and to determine that the change in discrimination exceeds a specified level. Accuracy is assessed by means of a tolerance test; change in discrimination is evaluated by partial F ratios. The tolerance of a variable not in an equation is one minus the squared multiple correlation between that variable and the other variables already in the equation. When the variable being tested is highly correlated with those already in the equation, its tolerance will be very small or near zero and inclusion of this variable in the computation of the inverse of the W matrix would cause inaccuracies. In Table 5.14 the tolerances are given for the variables not in the analysis after a particular step. For example, after the variable "aspect" is entered in step 1, the variable "river distance to rank 2 stream" (dstrnk2) has a tolerance of 0.50, indicating that it is somewhat correlated with aspect. On the other hand, elevation above water (elvwat) has a tolerance of 0.99, indicating that of the remaining variables not in the equation, it is the least correlated with aspect.

F-to-enter and *F-to-remove* are partial multivariate F statistics. F-to-enter does just what the name implies: it tests the significance of the additional discrimination that will be provided by a variable being considered for inclusion in the model. A small F value indicates that the variable will not offer much additional discriminating power. In Table 5.14, with no variables in the analysis, aspect has the highest F-to-enter, 7.2682; hence it is entered at step 1. Of the variables not in the analysis after step 1, distance to water (dstwtr) has the highest F-to-enter (6.5958), and it is the variable entered at step 2, followed in a similar manner by elevation above water (elvwat) at step 3. After step 3 none of the variables meet the minimum F-to-enter value of 4.0 used in the analysis. F-to-remove tests the significance of the decrease in discrimination that will occur should the variable in question be removed from those already selected. It is used to determine if there are any variables that no longer make a significant contribution to the discrimination. The variable with the greatest F-to-remove statistic makes the greatest contribution to discrimination. The second-largest F-to-remove value indicates the second most important variable, and so on. Of the variables included in the analysis at step 3, aspect has the largest F-to-remove (14.7298) and is deemed the most important variable, followed by dstwtr (13.0481) and elvwat (6.1746).

TABLE 5.14.

Stepwise output from Booth Mountain discriminant function analysis

VARIABLES NOT IN THE ANALYSIS AFTER STEP 0

Variable	*Tolerance*	*Minimum Tolerance*	*F to Enter*	*Wilks's Lambda*
RELIEF	1.0000000	1.0000000	1.3507	.96384
ASPECT	1.0000000	1.0000000	7.2682	.83202
DSTWTR	1.0000000	1.0000000	2.7678	.92861
DSTRNK2	1.0000000	1.0000000	.4695	.98713
ELVWAT	1.0000000	1.0000000	.5199	.98576
VIEW	1.0000000	1.0000000	3.4661	.91218

****

AT STEP 1, ASPECT WAS INCLUDED IN THE ANALYSIS

		Degrees of Freedom			*Significance Between Groups*
Wilks's Lambda	0.8320205	1	1	36.0	
Equivalent F	7.268164		1	36.0	.0106

VARIABLES IN THE ANALYSIS AFTER STEP 1

Variable	*Tolerance*	*F To Remove*
ASPECT	1.0000000	7.2682

VARIABLES NOT IN THE ANALYSIS AFTER STEP 1

Variable	*Tolerance*	*Minimum Tolerance*	*F To Enter*	*Wilks's Lambda*
RELIEF	.9962583	.9962583	1.4300	.79936
DSTWTR	.8646998	.8646998	6.5958	.70009
DSTRNK2	.5048064	.5048064	2.3535	.77960
ELVWAT	.9991313	.9991313	.5188	.81987
VIEW	.9817724	.9817724	4.0816	.74513

F STATISTICS AND SIGNIFICANCES BETWEEN PAIRS OF GROUPS AFTER STEP 1
(EACH F STATISTIC HAS 1 AND 36.0 DEGREES OF FREEDOM)

	GROUP 0
GROUP 1	7.2682 .0106

****

AT STEP 2, DSTWTR WAS INCLUDED IN THE ANALYSIS

		Degrees of Freedom			*Significance Between Groups*
Wilks's Lambda	0.7000884	2	1	36.0	
Equivalent F	7.496845		2	35.0	.0020

VARIABLES IN THE ANALYSIS AFTER STEP 2

Variable	*Tolerance*	*F To Remove*	*Wilks's Lambda*
ASPECT	.8646998	11.4245	.92861
DSTWTR	.8646998	6.5958	.83202

TABLE 5.14. Continued

VARIABLES NOT IN THE ANALYSIS AFTER STEP 2

Variable	*Tolerance*	*Minimum Tolerance*	*F To Enter*	*Wilk's Lambda*
RELIEF	.9879944	.8575272	.7627	.68473
DSTRNK2	.5044356	.4723641	1.7544	.66574
ELVWAT	.6555658	.5673605	6.1746	.59249
VIEW	.9748621	.8436076	4.1148	.62451

F STATISTICS AND SIGNIFICANCES BETWEEN PAIRS OF GROUPS AFTER STEP 2
(EACH F STATISTIC HAS 2 AND 35.0 DEGREES OF FREEDOM)

	GROUP 0
GROUP	
1	7.4968
	.0020

****

AT STEP 3, ELVWAT WAS INCLUDED IN THE ANALYSIS

		Degrees of Freedom			*Significance Between Groups*
Wilks's Lambda	.5924894	3	1	36.0	
Equivalent F	7.794996		3	34.0	.0004

VARIABLES IN THE ANALYSIS AFTER STEP 3

Variable	*Tolerance*	*F To Remove*	*Wilks's Lambda*
ASPECT	.7932439	14.7298	.84917
DSTWTR	.5673605	13.0481	.81987
ELVWAT	.6555658	6.1746	.70009

VARIABLES NOT IN THE ANALYSIS AFTER STEP 3

Variable	*Tolerance*	*Minimum Tolerance*	*F To Enter*	*Wilks's Lambda*
RELIEF	.9801711	.5583875	.3522	.58623
DSTRNK2	.5040043	.4538064	1.2885	.57023
VIEW	.9279786	.5651945	1.8992	.56025

F STATISTICS AND SIGNIFICANCES BETWEEN PAIRS OF GROUPS AFTER STEP 3
(EACH F STATISTIC HAS 3 AND 34.0 DEGREES OF FREEDOM)

	GROUP 0
GROUP	
1	7.7950
	.0004

MAXIMUM STEP REACHED

****

SUMMARY TABLE

	Action				
Step	*Entered*	*Removed*	*Variables In*	*Wilks's Lambda*	*Significance*
1	ASPECT		1	.832021	.0106
2	DSTWTR		2	.700088	.0020
3	ELVWAT		3	.592489	.0004

All of the stepwise procedures described here need to employ some type of stopping criterion to determine when maximum group separation has been parsimoniously achieved. Wilks's lambda is one popular criterion that takes into account the variability both within and between the groups. For example, in the summary table in Table 5.14 Wilks's lambda changes from 0.83 to 0.70 to 0.59 from step 1 to step 3. If a stopping criterion of 0.59 had been chosen, no additional variables would enter the analysis because their associated lambda values are less than the minimum. Lambda can also be converted into an overall multivariate F-ratio; in this situation the largest F value would be selected. The overall F values after steps 1, 2, and 3 are, respectively, 7.2682, 7.4968, and 7.7950 (Table 5.14). Alternatively, the partial F-to-enter can be employed. Any of the three tests will yield the same result.

Rao's V (Rao 1952:257), a generalized distance measure of the separation of the group centroids, can be used as a stopping criterion. A variable selected with this criterion may actually increase the within-group variance while adding to the overall group separation. The V statistic is calculated as

$$V = (n. - g) \sum_{i=1}^{p'} \sum_{j=1}^{p'} a_{ij} \sum_{k=1}^{g} n_j (X_{ik\cdot} - X_{i\cdot\cdot})(X_{jk\cdot} - X_{j\cdot\cdot})$$

and it measures the distance from each group centroid to the grand centroid weighted by the group size. That is, larger groups are more important than smaller ones in calculating the grand centroid. P' is the number of variables entered, including the current one. With large samples the distribution of V approximates that of chi-square with $p'(g\text{-}1)$ degrees of freedom. The change in V owing to the addition or deletion of variables can also be tested for statistical significance because it has a chi-square distribution with degrees of freedom equal to $(g\text{-}1)$ times the number of variables added or deleted. A variable should not be included if the change is not significant or if it is negative, indicating a decrease in group separation.

Other statistics based on the Mahalanobis squared distance between group centroids can be used to choose the variable that generates the greatest separation for the groups that are closest at that step. D^2, which gives equal weight to each pair of groups, is one of the statistics that can be used to measure group separation. The differences between two groups can also be tested with an F statistic, which gives greater weight to comparisons between large groups.

$$F = \frac{(n. - g - p' + 1)\, n_i n_j}{p'(n. - g)(n_i - n_j)} \quad D^2 (G_i \,|\, G_j)$$

Another criterion minimizes the residual variance:

$$R = \sum_{i=1}^{g-1} \sum_{j=i+1}^{g} \frac{4}{4 + D^2 (G_i \,|\, G_j)}$$

The average residual variance between groups can be obtained by dividing R by the number of pairs, or pairs can be weighted if some are given greater importance (Dixon 1973:243). R considers all pairs and hence tends to promote an equal separation of the groups.

MODEL VALIDATION AND GENERALIZATION

Validation

For predictive modeling purposes the best guide as to how well the statistical operations of multiple regression, logistic regression, and discriminant function analysis work is their performance on independent data. The accuracy and precision of sample-based predictions can be tested on independent data, on part of the sample that was excluded from the model-building process, or on internal criteria (see Chapters 7 and 8). In the validation procedure the coefficients of the calibration equation (model) are applied to independent data, and predicted values are generated that can be checked against actual values. The actual and predicted values can be compared on several levels of measurement and with a variety of statistical techniques, depending on the particular model. Kohler and Parker (1986) note that many of the predictive models presented in the archaeological literature during the past several years have not been validated and that there is little agreement about the "best" or "most correct" manner in which to undertake verification. When tried on fresh data, many procedures fail dismally. Failure of a predictive model may result from incorrect specification of the model, from the inclusion of too many parameters, or from sampling problems.

There are two levels of validation, simple and double. Simple validation involves testing a procedure on data different from those used to choose its coefficients. Double validation consists of testing the procedure (i.e., the particular equation) on data different not only from those used to choose its numerical coefficients but from those used to guide the choice of its form as well. By the *form* of the procedure we mean such questions as which variables of a total set enter an equation, whether the original or transformed variables are employed, whether products or ratios are used, and so on. Simple validation is more common since the double approach is usually prohibitively expensive in terms of time, money, and other considerations. To achieve as full a validation as possible, whether simple or double, the validation data should reflect the full range of variability expected in the final application.

Mosteller and Tukey (1977:37) note that when a predictive model is tested on the data that produced it overestimation of its performance is almost certain. This occurs because the optimization process that selects coefficients capitalizes on any and all idiosyncracies of a particular data set. As a result, the model will generally perform better for the data used in its development than for almost any other data

that will be encountered in practice. The strongest validation of a predictive model requires the utilization of data that were not included in its development.

Double Validation (Independent Data Procedures)

Double validation involves totally new data, possibly those gathered by another investigator after the form and coefficients of the procedure were determined. The most rigorous double-validation procedure involves additional field survey. Predictions can be generated for unsurveyed areas using some type of generalization procedure (reviewed below). The values of the independent variables must be obtained and then applied to the coefficients of the model (such as a classification function), and predicted values for the unsurveyed area must be generated. After the additional field survey is performed, the accuracy and precision of the predicted values of the dependent variable(s) can be assessed with the data values. If the goodness of fit between the predicted and observed values is poor, the original model may need to be reevaluated and the cause of the discrepancy determined.

Simple Validation (Split-Sample Procedures)

Whether a predictive modeler is stuck with simple validation or just content with it, the ease with which computers can handle repeated tasks makes several approaches feasible. Some validation techniques have been used more than others, and additional techniques remain to be explored. The basic idea is to divide the complete sample data set into subsets by some rational criterion or criteria and then to use one part to build the predictive model and the other to validate it. There is no best or unique answer to the question of how to subdivide a data set; several possibilities are outlined below.

In what some call the classical, half-and-half, or split-sample approach, the data are divided into two sets. One part is used to calibrate or build the model and the other is used for validation; sometimes, the two sets are interchanged and the process is repeated. It is possible to obtain a great deal of information about a body of data using a half-and-half approach, but many researchers do not have large enough samples to be able to do this and maintain any semblance of quality. Another problem is that the data cases may need to be randomized before splitting them into two sets to avoid trends or particular configurations among variables that may relate to the manner in which the original data were obtained.

Snee (1977) has considered the problem of choosing the one-half subset for model building and discusses the DUPLEX algorithm of R. W. Kennard. Basically, this algorithm tries to guarantee that the properties of the determinant of the covariance matrices are similar for the model calibration and verification sets. Snee also recommends that the data set not be split in half unless the total sample size is greater than $2p+25$, where p is the number of parameters in the model.

McCarthy (1976) suggests the use of more than one method of halving the data to obtain additional information. Instead of dividing the data into two parts, suppose it was divided into 10 parts. The model could then be calibrated on nine parts and verified on the tenth, and the process could be repeated nine times, verifying on a different part each time. By doing this, all of the data are used to assess the quality of what is to be gained by calibrating on a body of data nine-tenths of the total sample size. These procedures come closer than others to helping us to determine what level of performance can be expected from a predictive model.

Taking the process of multiple subsample validation to its logical end, one could calibrate the model on all but one case and then validate on the remaining case. This process is then repeated for every case. This selection procedure, referred to here as a jackknife test, but also called PRESS, was proposed by Allen (1971). It is in some respects a combination of all possible regressions, residual analyses, and validation techniques. If it were actually necessary to go through the computational process each time, this calibration-validation scheme could be prohibitive. Fortunately, it is often possible to calculate, either exactly or to a reasonable approximation, the effect of dropping an individual case or a small subset of the data. Draper and Smith (1981:326–327) feel that the PRESS method is advantageous because it provides detailed information about the stability of the model parameters over the sample space and because it can aid in defining influential data points.

Many statistical packages now calculate statistics that define the influence of individual cases. For example, one of the regression programs in the BMDP package calculates the residual that would be obtained if the case were omitted from the computation of the equation; the predicted value that would be obtained if the case were omitted; the Mahalanobis distance, which is the distance of the case from the mean of all cases used to estimate the regression equation; and Cook's distance (Cook 1979:15–18, 1979:169–174), a measure of the change in the coefficients of the regression that would occur if the case were omitted from the computations of the coefficient.

The results of a jackknife classification of the Booth Mountain nonsites (Group A) and sites (Group B) are given in Table 5.15. Incorrect classifications are noted by the letter of the group the case was placed in. The jackknifed Mahalanobis D^2 from the group centroid and the posterior probability $P(X|G)$ that the case (X) is a member of the indicated group (G) are also given. It should be noted that the probabilities are close, but not identical, to those presented in Table 5.13, which are not based on a jackknife classification. The jackknifed classification results are more conservative and realistic, especially for small sample sizes. Jackknife classification results are also more like the results that would be achieved with a set of independent data. Overall, 76.3 percent of the cases were classified correctly in the jackknife classification. Individually, 76.2 percent of the nonsites and 76.5 percent of the sites were classified correctly.

In the jackknife classification, four sites were misclassified as nonsites. Two of these, 5Pe746 and 5Pe749, were apparently misclassified because they are located 230 and 240 m, respectively, from the nearest water source. They are located near each

TABLE 5.15.

Jackknife classification of locations from Booth Mountain; nonsites are group A, sites are group B

	Incorrect Classifications	*Jackknifed Mahalanobis D^2 from and Posterior Probability for*			
		Group A		*Group B*	
GROUP A					
CASE					
1		21.7	.892	25.6	.108
2	B	19.7	.003	8.1	.997
3		1.3	.592	1.7	.408
4		.4	.846	3.5	.154
5		2.5	.726	4.1	.274
6	B	2.0	.471	1.4	.529
7		13.9	.998	25.8	.002
8		5.8	.991	15.0	.009
9	B	2.0	.500	1.7	.500
10		1.9	.854	5.1	.146
11		.8	.739	2.6	.261
12		1.3	.609	1.9	.391
13		3.4	.853	6.6	.147
14		3.3	.986	11.6	.014
15	B	1.5	.499	1.2	.501
16	B	8.4	.084	3.3	.916
17		1.0	.796	3.4	.204
18		1.2	.883	4.9	.117
19		2.1	.789	4.4	.211
20		1.2	.872	4.8	.128
21		1.1	.950	6.7	.050
GROUP B					
CASE					
22		8.2	.049	1.9	.951
23		3.6	.190	.4	.810
24		4.2	.256	1.7	.744
25		4.3	.163	.7	.837
26	A	1.3	.891	5.2	.109
27		2.5	.340	.9	.660
29		6.4	.161	2.7	.839
30		7.6	.067	2.0	.933
31		5.3	.208	2.3	.792
32		9.1	.100	4.3	.900
33		3.8	.484	3.3	.516
34	A	2.3	.669	3.4	.331
35	A	17.2	.785	19.5	.215
36		4.7	.172	1.2	.828
37		2.3	.283	.2	.717
38		4.9	.110	.4	.890
39	A	1.3	.594	1.7	.406

Note: Case 28 is a historical site that was not included in the analysis.

other and in association with several other sites that seem to date to the protohistoric period. 5Pe749 is a small lithic scatter that could be associated with 5Pe750, a large stone circle that is located about 200 m to the west, or with 5Pe746, which is located about 250 m to the north. The latter site consists of three standing wickiups and a small amount of scattered lithic artifacts. On the site form for 5Pe746 the distance to water is listed as 400 m, although the distance measured from the 7.5-minute quadrangle is only 230 m. Even so, the site is substantially farther from water than most others, which is surprising given the evidence of habitation. 5Pe746 is a good example of the fact that even when predictive models work, there is always the chance that an anomalous site will be misclassified.

The remaining two misclassified sites, 5Pe366 and 5Pe741, are both small rockshelters. Their misclassification points out the difficulty in predicting rockshelter sites. Clearly, more research is needed in the development of variables that can measure a locality's "rockshelter potential."

Even with all of the benefits noted above, simple validation is weaker, by some unknown amount, than it might initially appear to be. This is because the validation sample is often much more like the calibration sample than the target population to which one wishes to generalize. Frequently the calibration and validation data sets are obtained in the same sampling process. In most archaeological predictive modeling cases this appears to be the rule rather than the exception (see Chapter 7).

Simple Validation (Procedures with Simulated Data Sets)

Many predictive modeling studies use reclassification of the cases on which the procedure was developed as a measure of discriminatory effectiveness. Simple reclassification of the original cases will be biased and the efficacy of the technique inflated (this is known as upward bias). Many times, though, sample sizes are small, and too few cases can be withheld and used as verification data. One way to determine the bias involved is with simulated data. Frank et al. (1965) suggest two validation procedures that involve simulation. The first approach involves creating a synthetic random data set that bears no resemblance to the actual data on which the classification functions are based. The random data set can be designed to match the actual data set in terms of sample size, the number of groups, and variables. When the synthetic cases are classified by the function, any discriminatory power found can be interpreted as bias.

The second approach uses the original raw data. The cases in the original data matrix are randomized and then randomly assigned to groups, and then a discriminant analysis is performed. This process is repeated a number of times, and the results are averaged and used to assess the amount of bias present. In this approach the covariance matrices of the simulations will be similar to those of the original analysis. Berry (1984:843–853) has recently employed both of these techniques in a discussion of sampling and predictive modeling on federal lands in the western United States.

The validation procedures described above can be used with multiple regression and its many variants, with logistic regression, and with discriminant function analysis. Some are now standard options on computerized statistical analysis packages, such as BMDP, SAS, and SPSS, or they can be created easily by using sampling procedures to select the data to be withheld. These validation procedures, especially those applicable to smaller samples, need to be explored and exploited further because the "acid test" of any prediction scheme is its performance on independent data. Kohler and Parker (1986) hypothesize that the lack of validation in many predictive models occurs because it was not specified in a contractual obligation. This could be a healthy sign if agencies intend to have validation done by another party. Alternatively, it may indicate that the importance of validation has not been recognized as an essential ingredient of predictive modeling by the sponsoring agencies.

Generalization

Once a predictive model has been established and validated, the next step is to generalize the results to the target population. Essentially, in this process the results of a sample-based statistical procedure are employed to make inferences about a population. The predictions are based on what occurred at a sample of points or in a sample of quadrats, e.g., whether a site was or was not present or whether a site of a specific functional and temporal category was found. To be useful as a management tool these generalizations must be summarized; frequently this is done cartographically. While there is no limit to the data that can be represented on a map, the extrapolation of point or quadrat information to larger areas is not a simple task (see Chapter 10).

In order for the results of the sample-based procedure to be generalized to a larger area, it must be possible to extract the values for the independent or predictor variable from existing maps, whether manually or via some type of automated geographic system. This is not a trivial matter, because the lack of preexisting maps or of information that can be converted to a maplike format is the most commonly encountered problem in generating model-based predictions. The scale of resolution of the generalizations that can be produced is determined by the quality of the existing information. If one of the discriminating variables in a predictive model is only available on a section-by-section basis, generalizations can only be produced on that scale. Alternatively, if the required information exists at a very fine level of resolution for an area, it may be possible to make inferences at that scale. A final problem with secondary information extracted from maps is that maps are also simplifications of reality; the mere existence of the data is not an indication of their quality. The chances are that point or even quadrat values extracted from a map are probably themselves interpolated values, whether the variable is elevation, soil type, or vegetation. The quality of the data extracted from maps affects all aspects of the predictive model. If the data are of poor quality, then predictions derived from them will likely be poor. Means of assessing the quality of preexisting data and of compensating for variability in quality are discussed in Chapter 7.

Maps showing the distributions of such environmental variables as soil types, vegetation communities, slope, elevation, and the like are frequently combined in composite overlays that may or may not be associated with site locations. For example, a particular type of site may be found on southwest exposures of certain soil types in piñon-juniper forest within a specified elevation interval. Unless these sets of data are manipulated in a rigorous, objective manner (e.g., a game-theory approach), composite overlays merely indicate that certain types of sites tend to be associated with particular configurations of environmental variables without providing any specifics as to the nature of that association (see Chapters 3 and 4).

Multiple regression, logistic regression, and discriminant function analyses make composite overlays a more valuable managerial tool. These procedures provide objective delineations of the environmental variables that influence site location, and these results can replace traditional intuitive projections. Rather than just representing the intersection of numerous information sets, these procedures permit weights to be assigned to particular mappable variables that may be important determinants of site location. This capability then allows the researcher to associate probabilities with particular environmental features.

The process by which the predictive model information is portrayed cartographically is crucial to the interpretations that can be made. Unless the sample on which the predictive model is based is systematic, which would be most unfortunate from a sampling perspective, the values of the variable to be displayed (usually called z) in an x,y coordinate system are not neatly spaced on a regular grid pattern. Most samples procured in predictive modeling endeavors consist of data points or quadrats scattered irregularly over a region. Most mapping packages and graphic display procedures, however, require that a regular grid be created from irregularly spaced data. The basis for most generalization operations will be a grid of values that is a numerical representation defined by two x and y coordinates of the surface to be displayed. In predictive modeling these values might be the probability that a site is present.

The phenomena to be mapped, the z values, can be portrayed in several different ways. Linear interpolations can be made between grid nodes to locate the points where a specific contour line will cross the edge of a grid cell. The grid can also be drawn in perspective, with each grid node vertically offset by an amount proportional to the z value. This operation results in a block diagram and does not require interpolation. Unfortunately, it is more difficult to extract quantitative information from a block diagram than from a contour map. (See Chapters 8 and 10 for detailed discussions of mapping the results of predictive models.)

The predictive model for Booth Mountain, described above in the context of discriminant function analysis, can be presented as a contour map (Figure 5.12) generated with the SURFACE II GRAPHICS SYSTEM (Sampson 1978). SURFACE II is a software system for creating displays of spatially distributed data. The Booth Mountain map used only the 38 nonsite/site locations employed in the discriminant function model. For each location, easting and northing UTM coordinates were used for x and y scales. The third value was the cell's posterior probability of

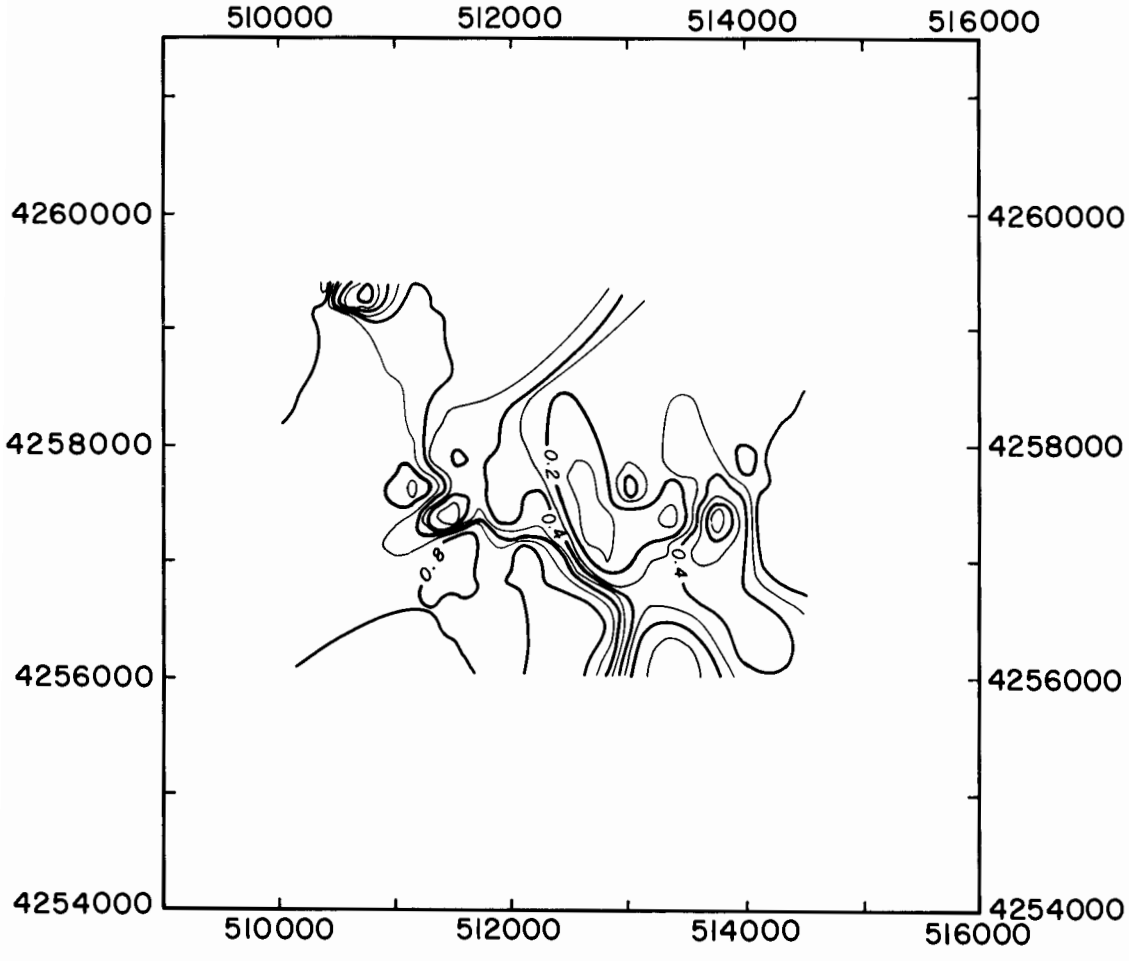


Figure 5.12. Posterior probability of site presence, Booth Mountain.

belonging to the site grouping (Table 5.15). The SURFACE program then interpolated between known points to create a generalized probability map, with probability scores ranging potentially from 0.0 to 1.0. The contour interval plotted was 0.1.

The Booth Mountain region was not as rigorously tested in the Fort Carson survey as the two major drainage basins, Red Creek and Turkey Creek (Altschul and Rose 1986). Because its importance was initially underestimated, only small portions of the mountain were surveyed. Furthermore, because the survey locations were not selected through a probabilistic sample design, it is difficult to generalize the results. Instead of generalizing the discriminant function model to the entire mountain, we felt only a portion of the area could be reliably modeled. The area modeled is a rectangular region encompassing the middle of the mountain, specifically between UTM northings 4256000 and 4259500 and eastings 13510000 and 13514500. About 50 percent of this area had been surveyed.

The areas of highest site probability on Booth Mountain lie on the southern and western slopes. Sites are found along small drainages that flow into Booth Gulch instead of directly into Turkey Creek. The entire eastern half of Booth Mountain appears to have been deliberately avoided by the prehistoric occupants. The only sites in the east are on the lower slopes of the mountain directly overlooking Turkey Creek (on Figure 5.12 this is the small high-probability zone located at about UTM northing 4257000 and easting 13514000). These sites tend to date to the Late Prehistoric period or earlier. In contrast, most sites on the western and southern slopes appear to date to the Protohistoric period. Therefore, it seems questionable whether the sites on the eastern slopes are culturally related to those located on the southern and western slopes.

We can only speculate about the reasons for the intensive occupation of Booth Mountain. It would have provided a relatively safe refuge for small groups for short periods of time. There is also no question that during the latter half of the nineteenth century bands of Indians were intermittently seeking such protection. One possibility is that the interior slopes of Booth Mountain were selected for occupation because of their inaccessibility.

The Booth Mountain predictive model remains to be tested with independent data, but it would seem that areas with p values greater than 0.5 have reasonably good chances of containing sites. Areas with p values less than 0.5 would still need to be treated cautiously until the model is tested, nor should the model be extended to other parts of Booth Mountain in the absence of additional survey data.

CONCLUSIONS

The topics discussed in this chapter cover a wide range of material, but they indicate that the construction of predictive models is a multistep process. The procedure that has been advocated emphasizes an understanding of the nature of

the basic data on which the predictive models are based. Controlling for temporal and functional variability in the data is critical because of the effects of such variability on the resolution of the predictive model. The basic data used in models are the measurements for variables that have theoretical importance but each of which represents only one dimension of variability. It was emphasized that the scales of these measurements must be congruent with the assumptions of the statistical model that will be employed. Distribution of the values is also important. Descriptive statistics can be used to determine if the variables are normally distributed; if they are not, alternative probability distributions can be used to determine the nature of the differences. Finally, bivariate statistical techniques can be employed to examine relationships between pairs of variables prior to their incorporation into a predictive model.

After this discussion of the basic data used in modeling, three different types of predictive models were discussed. These models are applicable to many different types of dependent and independent variables. Multiple regression is normally used when both the dependent and independent variables are measured on interval and ratio scales. When the dependent variable is categorical and the independent variables are measured on any scale (from nominal to ratio), logistic regression provides an acceptable alternative. If all of the independent variables are measured on an interval or ratio scale, however, discriminant function analysis may be more effective.

Regardless of the statistical procedure used to produce a predictive model, verification of the results is an important part of the modeling process. Several different procedures have been described that are applicable in different situations. Finally, some inherent problems in producing a graphic display to portray the results are discussed. Although the graphic example used provides a relatively simple model when compared with the results obtained from geographic information systems, it does highlight the steps involved in computing these models, which are often lost in the inner workings of the computer.

In addition to the people mentioned in the acknowledgments for Chapter 3. Martin Rose would like to thank his coworkers at Statistical Research, Inc., under whose auspices this chapter was written. The comments, queries, and clarifications made by June-el Piper, Lynne Sebastian, Mike Garratt, Dan Martin, Jim Judge, Ken Kvamme, and the other volume authors and reviewers of the draft chapter and its subsequent revisions were particularly helpful. The authors greatly appreciate all of their time and effort, without which the current status of the manuscript would have been impossible.

REFERENCES CITED

Allen, D. M.
1971 Mean Square Error of Prediction as a Criterion for Selecting Variables. *Technometrics* 13:469–475.

Alexander, Robert K., John D. Hantley, and Thomas Babcock
1982 *A Settlement Survey of the Fort Carson Military Reservation. Grand River Consultants, Grand Junction, Colorado.* Submitted to Interagency Archeological Services, Contract No. C3572[78]. Copies available from Interagency Archeological Services, Denver.

Altschul, Jeffrey H., and Martin R. Rose
1986 Predictive Models of Prehistoric Site Location. In Historic Preservation Plan for Fort Carson Military Reservation, Colorado, assembled by Christian Zier. Centennial Archaeology, Inc. Draft report submitted to Interagency Archeological Services, Contract No. CX-1200-5-A00G. Copies available from Interagency Archeological Services, Denver.

Berry, Michael S.
1984 Sampling and Predictive Modeling on Federal Lands in the West. *American Antiquity* 49:842–853.

Blalock, Hubert M., Jr.
1972 *Social Statistics.* McGraw-Hill, New York.

Carmichael, J. W., and P. H. A. Sneath
1969 Taxometric Maps. *Systematic Zoology* 18:402–415.

Carter, G. M., and J. E. Rolph
1974 Empirical Bayes Methods Applied to Estimating Fire Alarm Probabilities. *Journal of the American Statistical Association* 69:880–885.

Chernoff, Miriam
1982 Empirical Bayes Estimation of Ceramic Proportions at Teotihuacan. Paper presented at the 47th Annual Meeting of the Society for American Archaeology, Minneapolis.

Clark, Isobel
1979 *Practical Geostatistics.* Applied Science, London.

Cliff, A. D., and J. K. Ord
1973 *Spatial Autocorrelation.* Pion, London.

Cook, R. D.
1979 Detection of Influential Observations in Linear Regression. *Technometrics* 19:15–18.

Cooley, William W., and Paul R. Lohnes
1971 *Multivariate Data Analysis.* John Wiley and Sons, New York.

Cowgill, George L.
1975 A Selection of Samplers: Comments on Archaeo-Statistics. In *Sampling in Archaeology*, edited by James Mueller, pp. 258-274. University of Arizona Press, Tucson.

Daultrey, S.
1976 *Principal Components Analysis.* Catmog 8. Concepts and Techniques in Modern Geography. Geo Abstracts, University of East Anglia, Norwich, England.

Derman, Cyrus, Leon Gleser, and Ingram Olkin
1973 *A Guide to Probability Theory Application.* Holt, Rinehart and Winston, New York.

Dixon, W. J., editor
1981 *BMDP Statistical Software 1981.* University of California Press, Berkeley.

Draper, Norman R., and Willard E. Lawrence
1970 *Probability: An Introductory Course.* Markham, Chicago.

Draper, N. R., and H. Smith
1970 *Applied Regression Analysis.* John Wiley and Sons, New York.

1981 *Applied Regression Analysis*, 2nd ed. John Wiley and Sons, New York.

Efron, B., and C. Morris
1973 Stein's Estimation Rule and Its Competitors—An Empirical Bayes Approach. *Journal of the American Statistical Association* 68:117–130.

1975 Data Analysis Using Stein's Estimator and Its Generalizations. *Journal of the American Statistical Association* 70:311–319.

1977 Stein's Paradox in Statistics. *Scientific American* 236(5):119–127.

Everitt, Brian
1974 *Cluster Analysis.* Heinemann Educational Books, London.

Fisher, W. D.
1936 Use of Multiple Measurements in Taxonomic Problems. *Annals of Eugenics* 7:179–188. London.

Frank, Ronald E., William F. Massy, and Donald G. Morrison
1965 Bias in Multiple Discriminant Analysis. *Journal of Marketing Research* 2:250-258.

Gower, J. C.
1971 A General Coefficient of Similarity and Some of Its Properties. *Biometrics* 27:857–872.

Grady, James
1980 *Environmental Factors in Archaeological Site Locations, Piceance Basin, Colorado.* Cultural Resource Series No. 9. Bureau of Land Management, Denver.

Grizzle, J. E., C. F. Starmer, and G. G. Koch
1969 Analysis of Categorical Data by Linear Models. *Biometrics* 25:489–504.

Gunst, Richard F., and Robert L. Mason
1984 *Regression Analysis and its Application: A Data Oriented Approach.* M. Dekher, New York.

Harris, Bernard
1966 *Theory of Probability.* Addison-Wesley, Reading, Massachusetts.

Harris, Richard J.
1975 *A Primer of Multivariate Statistics.* Academic Press, New York.

Hartwig, Frederick, and Brian E. Dearing
1979 *Exploratory Data Analysis.* Quantitative Applications in the Social Sciences Paper No. 16. Sage Publications, Beverly Hills.

Hayes, William H., and Robert L. Winkler
1971 *Statistics: Probability, Inference, and Decision.* Holt, Rinehart and Winston, New York.

Hodder, Ian, and Clive Orton
1976 *Spatial Analysis in Archaeology.* Cambridge University Press, Cambridge, England.

Jennich, R. I., and R. H. Moore
1975 Maximum Likelihood Estimation by Means of Nonlinear Least Squares. In *Proceedings of the Statistical Computing Section*, pp. 57–65. American Statistical Association, New York.

Johnston, V.
1972 *Econometric Methods*, 2nd ed. McGraw-Hill, New York.

King, L. V.
1969 *Statistical Analysis in Geography.* Prentice-Hall, Englewood Cliffs, New Jersey.

Kintigh, Keith
1984 Measuring Archaeological Diversity by Comparison with Simulated Assemblages. *American Antiquity* 49:44–54.

Klecka, William
1980 *Discriminant Analysis.* Sage Publications, Beverly Hills.

Kohler, Timothy A., and Sandra C. Parker
1986 Predictive Models for Archaeological Resource Location. In *Advances in Archaeological Method and Theory*, vol. 9, edited by Michael Schiffer, pp. 397–452. Academic Press, New York.

Lapin, Lawrence
1978 *Statistics for Modern Business Decisions*, 2nd ed. Harcourt, Brace, Janovich, New York.

Larralde, S., and S. M. Chandler
1981 *Archaeological Inventory in the Seep Ridge Cultural Study Tract, Uintah County, Northeastern Utah.* Cultural Resource Series No. 11. Bureau of Land Management, Salt Lake City.

MacNeish, Robert S.
1964 Ancient Mesoamerican Society. *Science* 143:531–537.

Mahalanobis, P. C.
1936 On the Generalized Distance in Statistics. *Proceedings of the National Institute of Science* 12:49–55. Calcutta, India.

McCarthy, P. J.
1976 The Use of Balanced Half-Sample Replication in Cross-Validation Studies. *Journal of the American Statistical Association* 71:596–604.

Morrison, Donald F.
1976 *Multivariate Statistical Methods.* McGraw-Hill, New York.

Mosteller, F., and J. W. Tukey
1977 *Data Analysis and Regression.* Addison-Wesley, Reading, Massachusetts.

Nie, Norman H., C. Hadlai Hull, Jean G. Jenkins, Karin Steinbrenner, and Dale H. Bent
1975 *SPSS: Statistical Package for the Social Sciences,* 2nd ed. McGraw-Hill, New York.

Rao, C. R.
1952 *Advanced Statistical Methods in Biometric Research.* Wiley, New York.

Ripley, Brian D.
1981 *Spatial Statistics.* John Wiley and Sons, New York.

Rogers, A.
1974 *Statistical Analysis of Spatial Dispersion.* Pion, London.

Rogge, A. E., and T. Lincoln
1984 Predicting the Distribution of Archaeological Sites: A Case Study from the Central Arizona Project. Paper presented at the 49th Annual Meeting of the Society for American Archaeology, Portland.

Sampson, R. V.
1978 *Surface II Graphics System.* Kansas Geological Survey, Lawrence, Kansas.

Sanders, William T., Jeffrey Parsons, and Robert Santley
1979 *The Basin of Mexico.* Academic Press, New York.

Scholtz, S.
1981 Location Choice Models in Sparta. In *Settlement Predictions in Sparta,* by R. Lafferty, J. Otinger, S. Scholtz, W. F. Limp, B. Watkins, and R. Jones, pp. 207–222. Arkansas Archaeological Survey Research Series No. 14. Fayetteville.

Snedecor, George W., and William G. Cochran
1978 *Statistical Methods*, 6th ed. Iowa State University Press, Ames.

Snee, R. D.
1977 Validation of Regression Models: Methods and Examples. *Technometrics* 19:415–428.

Sokal, R. R., and P. H. A. Sneath
1963 *Principles of Numerical Taxonomy*. Freeman, London.

Stevens, S. S.
1946 On the Theory of Scales of Measurement. *Science* 103:677–680.

Tatsuoka, Maurice
1971 *Multivariate Analysis: Techniques for Educational and Psychological Research*. John Wiley and Sons, New York.

Thomas, David Hurst
1976 *Figuring Anthropology: First Principles of Probability and Statistics*. Holt, Rinehart and Winston, New York.

Thomas, R. W.
1977 *An Introduction to Quadrat Analysis*. Concepts and Techniques in Modern Geography No. 12. Geo Abstracts. University of East Anglia, Norwich, England.

Tukey, J. W.
1977 *Exploratory Data Analysis*. Addison-Wesley, Reading, Massachusetts.

Unwin, D. J.
1981 *Introductory Spatial Analysis*. Methuen, London.

Van de Geer, J. P.
1971 *Introduction to Multivariate Analysis for the Social Sciences*. W. H. Freeman, San Francisco.

Wallis, J. R., N. C. Matalas, and V. R. Slack
1974 "*Just A Moment*." Water Resources Research 10(2). American Geophysical Union.

Wood, John J.
1971 Fitting Discrete Probability Distributions to Prehistoric Settlement Patterns. In *The Distribution of Prehistoric Population Aggregates*, edited by George J. Gumerman, pp. 63–82. Prescott College Anthropological Reports No. 1. Prescott, Arizona.

Wrigley, N.
1976 *Introduction to the Use of Logit Models in Geography*. Concepts and Techniques in Modern Geography No. 7. Geo Abstracts. University of East Anglia, Norwich, England.

Chapter 6

COLLECTING NEW DATA FOR THE PURPOSE OF MODEL DEVELOPMENT

Jeffrey H. Altschul and Christopher L. Nagle

In Chapter 3 it was suggested that the first step in model building is to identify both short- and long-term objectives. Subsequently, existing data pertaining to site location—archaeological records and reports, ethnographic information, historical and ethnohistoric accounts, and macro- and microenvironmental data—can be collected to meet these objectives. An evaluation of the existing data (a topic discussed in detail in Chapter 7) will often reveal data gaps, which must then be filled through the collection of new data before the modeling process can continue.

If the data gaps are relatively minor, they can often be filled through limited surveys targeted to specific model requirements. Funding for research-oriented projects is limited, however, especially in cultural resource management contexts. An alternative source of data to fill specific gaps is the results of the inventory surveys that are routinely conducted on relatively small parcels of land to fulfill legal requirements. Fieldwork on these surveys is oriented toward meeting compliance criteria and not toward research, and as a result data from inventory surveys have often not been incorporated into regional plans, such as statewide comprehensive plans or regional predictive models. Given that most predictive models based on existing data need independent data against which to validate the model, the failure to utilize inventory survey results is not only an unfortunate decision but also in the long run an extremely expensive one.

Even with the inclusion of some new data there is a limit to the predictive power of a model based on existing data. As discussed in greater detail in Chapter 7, these models are primarily limited by the quality and representativeness of the existing data base. When the data are of uneven quality or are biased in favor of certain resources or areas, the degree to which the resulting model can be generalized is greatly restricted, and the use of more powerful modeling techniques is generally unwarranted. In such situations it may be prudent to start fresh and collect data that can be used to obtain the designated modeling objectives.

PLANNING FOR FIELDWORK

While individual research projects will have their own specific goals, all projects designed to gather data to be used in predictive models of site location must share three critical objectives. First, these projects must evaluate the range of variability in the nature of sites and site locations. Certain parameters of the phenomena being modeled must be determined, such as the number of site types (including how the term *site* should be defined), the covariation between site locations and environmental attributes, and the potential relationship among sites. Since these features are determined by generalizing the results of sample data to a larger area, the sample must be derived in an unbiased fashion, or at least the nature of the bias must be known in order to obtain usable estimates.

A second objective for all projects designed to gather data for predictive modeling should be to make certain that all "magnet" sites are found. From a predictive modeling standpoint, a magnet site can be defined as one whose location affects the location of other sites. While it may be argued that all sites affect the location of at least one other site (i.e., the site one moves to is related to the site one has just left), for the purposes of this discussion the term *magnet site* will be restricted to sites that affect settlement on a regional scale. In the absence of data about such sites, the model might yield a high percentage of wrong predictions. For instance, the location of a Hohokam agricultural hamlet was probably conditioned not only by environmental factors but also by the distance to the regional center with which it was affiliated. In these situations, predictive models based solely on environmental factors have not been particularly useful.

Magnet sites can not only exert a pull on other sites but can also repel them. The location of a regional center often appears to have been adjusted partly on the basis of the "pulls" and "pushes" of neighboring centers. As the above example indicates, magnet sites can be major centers like Spiro, Moundville, Cahokia, and Snaketown. But it is not just big, complex sites that act as magnets, nor are such sites confined solely to complex societies. The Archaic and Woodland periods in the western Ouachita Mountains of eastern Oklahoma provide an example of hunter-gatherer settlement systems affected by magnet sites. These settlement systems were characterized by an extremely stable pattern of winter basecamps in the river valleys and specialized activity camps dispersed throughout the region in the summer. While the special-activity camps were occupied on a temporary basis, many of the winter basecamps were occupied more-or-less continually for thousands of years (Altschul 1984). Further, evidence suggests that the locations of many special-activity sites covary with the location of the basecamps. Thus, while it would be possible to make generalized areal predictions of site location for this region, it would be useless to construct a point-specific predictive model without knowing the locations of the winter basecamps.

Magnet sites are culture-specific and thus change through time. One of the most commonly cited examples is Teotihuacan, a large precolumbian city located in the Basin of Mexico in a northeastern subvalley of the same name. Teotihuacan has

figured predominantly in discussions of sampling and survey strategies in archaeology for nearly two decades (Flannery 1976; Mayer-Oakes and Nash 1964; Schiffer et al. 1978). Time and time again it has been pointed out that if one implemented a 10 or 20 percent simple or stratified random sample survey design for the Teotihuacan Valley, there is a good chance that the entire city would be missed.

Knowledge of the location of Teotihuacan is critical to an understanding of settlement patterns in the Teotihuacan Valley, and indeed in the larger Basin of Mexico, from about 150 BC until about AD 750. It is not at all clear, however, that information about the city's location would be necessary for predicting settlement in the Teotihuacan Valley before 150 BC. Prior to the rapid growth of the city, settlement in this part of the Basin of Mexico was characterized by sites located primarily in the Patlachique Range overlooking the valley floor (Sanders 1965:92-93). In contrast, the city of Teotihuacan was located near freshwater springs on the valley floor. Inclusion of the city's location in a predictive model of pre-Christian era settlement in the region might well lead to inaccurate predictions.

An alternative explanation for the lack of evidence for early occupation of the area that was subsequently encompassed by the city of Teotihuacan is that such evidence was destroyed or buried by later construction, an alternative that brings us to the third objective. In order to gather data intended for use in predictive modeling, we must determine how behavior becomes part of the archaeological record and how postdepositional processes affect our ability to discover sites and to interpret them (Chapter 4). While many archaeologists have become increasingly interested in depositional and postdepositional processes, only rarely are these factors explicitly considered in survey design. Too often cultural resource management surveys make use of standardized field techniques that do not take the specific situation into account. Thus, archaeologists often use pedestrian surface survey in all portions of a target area, including heavily alluviated valleys and highly dissected bluffs. In the eastern woodlands of the United States, survey techniques often include transects of systematic shovel tests, but the interval between tests and the size and depth of the tests are often specified in the contract rather than being chosen with regard to geomorphic conditions or to the size and nature of the expected sites.

Even with the most appropriate techniques, however, it is highly unlikely that all sites in a survey tract will be found. If the probability of discovery for every site were equal and independent of all others, then at least the resulting parameter estimates would not be biased. But discovery probabilities are not equal. Smaller sites cover less surface area than larger ones and thus have less chance of being found. Special-activity sites in general have fewer remains than more permanent ones and therefore are less likely to be found.

The problems cited above have to do with visibility; thus, for sites that exhibit some surface manifestations the main objective is to equalize or at least control for differing discovery probabilities. A very different problem exists, however, for sites that exhibit no surface expression. Sites that have been buried by natural processes, such as river alluviation, mud flows, or shifting sands, or components that are

masked by later occupations are the most difficult to predict. Archaeologists have generally dealt with this problem by ignoring it. Lip service is usually paid to the Quaternary geology of an area, but rarely is the study of geomorphic land surfaces made an integral part of a predictive modeling project's research design. Without such studies, however, it is virtually impossible to construct accurate predictive models.

SURVEY STRATEGIES

In order to create a predictive model, a survey design must be developed that will provide sufficient data to calculate estimates on various aspects of sites and site locations, allow the identification of all or of a high proportion of magnet sites, and allow us to assess the effects of depositional and postdepositional processes. By nature, such a design must be multifaceted since each of these objectives can best be met through a different survey strategy. For example, parameter estimates rely on some type of probabilistic sampling foundation, while discovery of magnet sites or paleo land surfaces is best done by purposely selecting areas for examination. In the following sections, appropriate survey strategies to meet each of the three objectives are discussed.

Probabilistic Selection

One of the main goals of a reconnaissance survey program is to obtain reliable estimates on a variety of site attributes in a region from only studying a portion of that region. Usually, statistical inferences are drawn concerning site density, the proportion of different site types, the covariation between site locations and environmental attributes, and so on. These are basically ideographic or descriptive observations about the population. The problem, then, is how to select a sample of cases from which descriptive observations about the population can be inferred with a reasonable degree of confidence.

Much of probability sampling theory is concerned with this issue. It is important to point out that there is nothing in probability theory that guarantees the "correctness" of sample estimates. Rather, the advantage of using probability theory to select samples is that it allows us to control the bias in the selection procedure. As Cowgill notes, "the advantage of probability sampling is not that it very often enables us to be *sure* about the population, but that it helps us . . . to *know where we stand* in relation to various inferences" (1975:262, emphasis original).

The literature on sampling in archaeology written during the last two decades is voluminous (Hodder and Orton 1976; Mueller, ed. 1975; Nance 1983; Orton 1980; S. Plog 1976, 1978; S. Plog et al. 1978; Redman 1974; Sanders et al. 1979). It is not our purpose here to review this literature or to restate the basic principles of sampling. Instead, we wish to introduce some of the issues that should be considered in the design of a probabilistic sample survey.

In predictive modeling the primary objects of interest are sites (however defined) and their locations. In the abstract we might consider all the sites in a region as the population of interest. Sampling then would consist of selecting a number of these sites for observation. In reality, of course, if we knew where all the sites were located we would have no need for a predictive model of site location.

In practice we do not know where all the sites are or even what proportion of the total number of sites has previously been recorded. The most common approach in this situation is to use the region as the frame of reference. The region can then be divided into a number of smaller units such that all portions of the region are situated in one and only one unit. These areas then become the sampling units.

The use of spatial units as samples from which to make inferences about cultural phenomena has led to a certain amount of confusion among archaeologists (Mueller 1975) that only recently has begun to be resolved (Nance 1983; S. Plog et al. 1978). At issue is the difference between element and cluster sampling. Simply put, element sampling requires that each element of the population be considered a distinct sampling unit. If the sampling unit is a specified area, and if our interest focuses on attributes of those units, then we are conducting element sampling. Common examples in archaeology include estimates of the mean number of sites per grid unit or of the proportion of a site type per grid unit. Here we are conducting element sampling because the elements of interest are the grid units, *not* the sites themselves. In contrast, if our interest focuses on making inferences about attributes of sites found in the grid units then we are engaged in a form of cluster sampling. The most common use of cluster sampling in predictive modeling is in point-specific models (e.g., Kvamme 1983; Larralde and Chandler 1981; Reed and Chandler 1984). In these models, all sites found in the various grid units are combined into one group whose environmental attributes are compared with the attributes of a group of nonsite locations; the results are then generalized to the survey universe.

The failure to distinguish between cluster and element sampling leads not just to confusion but to miscalculation of basic statistics. The most common error is to calculate parameter estimates as if the cluster sample data had been collected by element sampling. In general this will lead to underestimation of the sampling error. The obvious solution is to be clear about the type of sample one is working with and then to use the correct equations to calculate the estimates. Unfortunately, we usually design surveys with a multitude of objectives. During analysis of the survey results, then, the object of interest, or the sampling element, varies, as does its relationship to the sampling unit. Although we will have some idea of the types of sampling elements in which we will be interested before the survey and analysis are carried out, we cannot foresee all the possible units of interest. This is one of the main reasons why flexible, multistep sampling designs are well suited to archaeological research.

The first task in designing a survey is to subdivide the region into sampling units and then to select a specified number of these units to survey. The immediate questions to answer are what size and shape should the units be and how many do

we need to survey. The critical factor in determining the precision of an estimate is the absolute number of units surveyed. In a 100,000 ha region, for example, given a known population variance, a sample of 1000 units, each 10 ha in size, will yield an estimate of site density that is twice as precise as that derived from a sample of 250 units, each 40 ha in size. This does not mean, however, that the best formula for selecting sample units is simply to divide the region into very small units; logistical cost and locational errors escalate as unit size decreases.

We have been discussing these issues primarily from the perspective of element sampling. In the site-density example given above, all the sampling units enter into the computation of the standard error. If our interest shifts from attributes of sampling units to attributes of sites, then the effective number of sampling units shifts from the total number of units surveyed to the number that contained the element of interest (in this case, sites). Thus, if we had surveyed 250 40 ha units and found a total of 30 sites distributed among 20 of those units, the effective number of sample elements would not be 250 (the number of units surveyed), nor would it be 30 (the number of sites). The effective number of elements would be 20, the number of sampling units that contained sites. Statisticians have suggested that, as a rule of thumb, 30 or more such units are needed before variance estimates and confidence intervals can be computed (e.g., Cochran 1977; Dixon and Leach 1978; see also Nance 1983:340). A perusal of the predictive modeling literature in archaeology indicates that this condition is rarely met.

The size and shape of the sample unit has been the subject of much discussion within archaeology (Judge et al. 1975; Mueller 1974; S. Plog 1976; S. Plog et al. 1978; Sanders et al. 1979; Schiffer and Wells 1982; Schiffer et al. 1978). Two types of sample units are common in archaeology, square quadrats and rectangular transects. There have been experiments with other types of units, such as circles (e.g., Goodyear 1975), but these are not generally used. Much of the discussion surrounding whether to use quadrats or transects concerns the so-called edge effect. In surveys, all sites found in the sample units are usually recorded. This includes sites that are contained wholly within the unit as well as those that are only partially located in the unit. Thus, when a unit is surveyed the crew actually surveys an area that is somewhat larger than the unit itself. How large? This depends on the average size of the site. We will illustrate this problem with an example drawn from a similar discussion in S. Plog et al. (1978:399). Let us assume that all sites in a region have a radius of 50 m. If we survey four square quadrats in that region that are 0.25 km² and are situated as in the upper part of Figure 6.1, an area of 1 km² is examined. But because all sites whose centers are located within 50 m of any border will be recorded, the survey will actually cover 1.44 km². To show the effect of sample unit size, assume that the same area had been surveyed in two rectangular quadrats, each of which was 0.50 km² (the lower portion of Figure 6.1). The surveys crews would still walk the same 1 km² area, but now they would only record sites in a 1.30 km² area. Finally, the edge effect differs between transects and quadrats. In the case proposed above, a quadrat with 500 m long sides actually covers 0.25 km² with a hypothetical coverage extending to an 0.36 km² area. In contrast, a transect 250 by

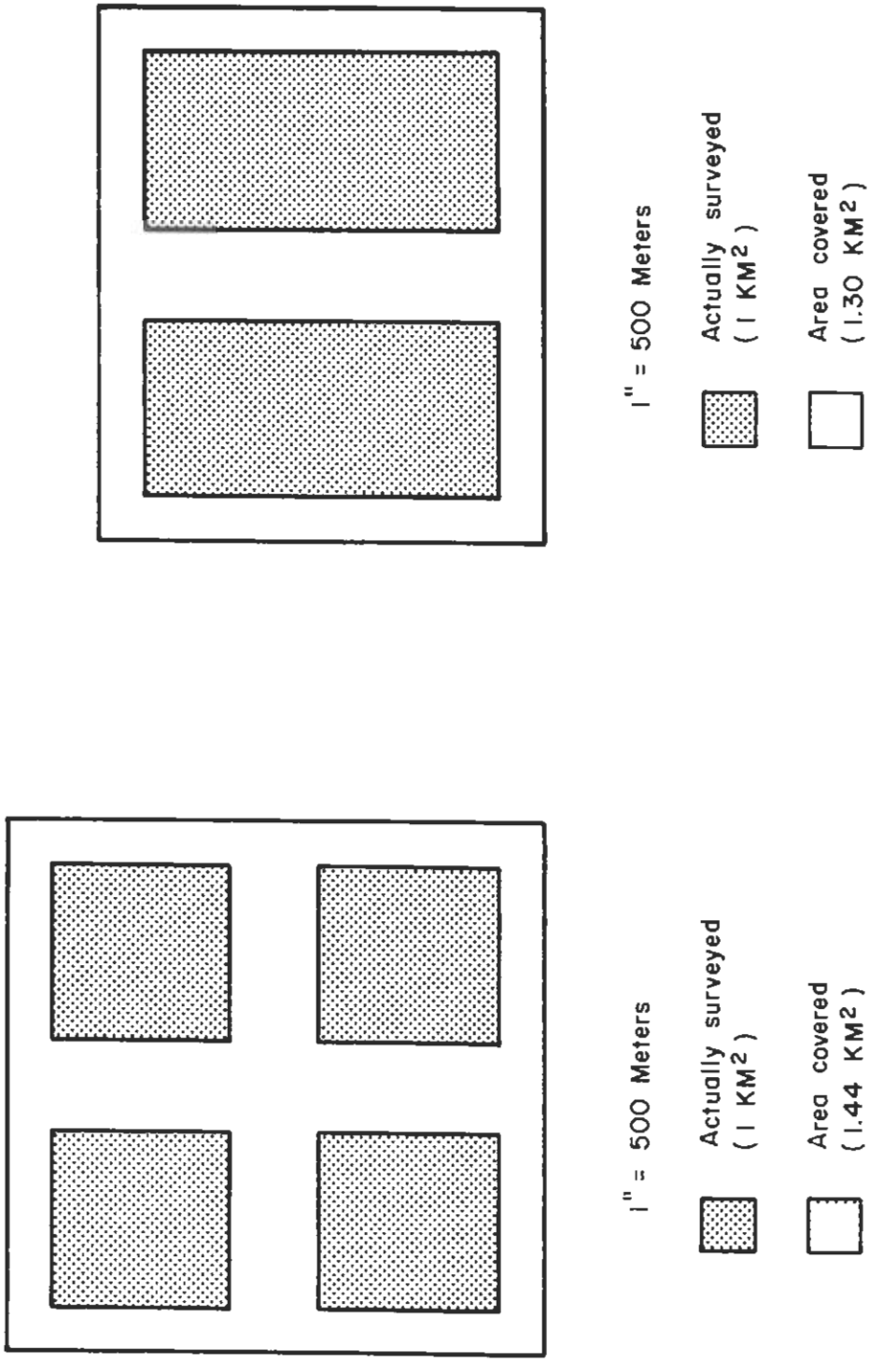


Figure 6.1. Edge effect for survey units with different sizes and shapes.

1000 m also covers an area of 0.25 km^2, but the hypothetical unit surveyed is 0.385 km^2. Thus, the edge effect is greater as the size of the sites increases, as the size of the sample unit decreases, and as the ratio of the length to width of the sample units increases.

The edge effect is in large part the reason why site density and the total number of sites tend to be overestimated. For example, if a 10 by 10 km (100 km^2) area is sampled using 40 square quadrats, each 500 m on a side, the crews would actually walk 10 km^2 or 10 percent of the survey universe. But if none of the units adjoined and all sites were 50 m in radius, then all sites in a 14.4 km^2 area should be found (i.e., $0.36\ km^2 \times 40 = 14.4\ km^2$). S. Plog et al. (1978:395–400) have analyzed this situation in some depth. Their work, drawn from simulations of survey results from the Valley of Oaxaca, indicates that the smaller the sampling fraction the greater the problem of overestimation. This is exactly what one would expect. As the sampling fraction increases, the likelihood of selecting contiguous survey units also increases, and the larger the number of contiguous units the smaller the edge effect.

Once the edge effect is understood, it can be compensated for in the calculation of parameter estimates. F. Plog (1981:32) has suggested that the hypothetical area covered by a survey unit be used in calculating site density rather than the actual unit area. The hypothetical area covered is found by determining the average site size and increasing the effective area of the unit by the radius of an average site. Thus, if the average site radius is 50 m then a 500 by 500 m unit would become a 600 by 600 m unit, effectively changing the area from 0.25 to 0.36 km^2. Nance (1983:311) has suggested using modal site size as a more appropriate measure because site size usually follows a highly skewed distribution.

As illustrated above, transects—with their greater edge effect—allow us to cover larger areas than quadrats. Thus, we can expect that more sites will be found in a transect survey than in a quadrat survey of the same amount of land. As long as steps are taken to compensate for the edge effect, there are several good reasons for using transects, especially when our objective is to construct an inductively based correlative model. Since the site is most often the subject of interest in these types of predictive models, and since the absolute number of sites is usually the determining factor of the power of the model, a transect design is the best choice for obtaining data for correlative purposes. This is especially true when we divide the total number of sites among various site types and then try to model the individual site types. Although in theory these models make the most sense (i.e., members of the individual groups all represent the same type of cultural behavior), in practice they are rarely created because the sample sizes of the individual site classes are almost always too small. In addition, although the evidence is weak, most studies done on archaeological data have indicated that transects yield more precise estimates than quadrats (see Judge et al. 1975; S. Plog 1976; Sanders et al. 1979; cf. Mueller 1974). Finally, from a logistical point of view, transects are generally easier to lay out in the field and are less prone to locational errors than are quadrats (Schiffer et al. 1978).

The comparison between survey units, however, is not all one-sided. Quadrats are probably more useful in studies of the relationship between sites and the environment. Precisely because of the smaller edge effect and lower length-width ratio of these units, it is easier to control for and characterize surrounding environmental features. Thus, when our interest shifts from obtaining data primarily on site location to obtaining data on how sites are situated in relation to environmental features (as is the case for many explanatory models), then quadrats may be the appropriate choice of sampling unit type.

The choice of size and shape for the sampling units depends on what we are trying to estimate and on making our best guess about the underlying site distribution. Most discussions about this issue assume that the parameter of interest will be either site density or the total number of sites (Judge et al. 1975; Matson and Lipe 1975; S. Plog 1976; S. Plog et al. 1978; Sanders et al. 1979). In cases where this assumption is correct, the primary factors to consider in choosing a sampling unit have to do with how these units affect the number of sites discoveted. During parts of the predictive modeling process, however, we are concerned with more than just numbers of sites. We must also determine the relationship between sites and the surrounding environment and between sites and other sites of the same site class and different site classes. To study these latter issues we need grid units that are large enough to capture these patterns, and we need to survey enough of these grid units so that we can generalize the resulting patterns with sufficient confidence.

The diverse objectives discussed above again strongly point up the need for a multistep survey strategy. For example, we might survey a large number of small grid units to obtain relatively precise estimates of site density and then use a smaller number of large grid units to determine the spatial relationship between sites. Whatever the exact strategies chosen, it is important to remember that several types of information are required in order to predict site locations. Use of a single, set strategy will probably mean that a larger proportion of the sample universe must be surveyed in order to obtain the same degree of precision on a number of estimates as can be derived from a series of sampling straregies, each focused on a separate target.

Once we have decided on a sampling unit, the next question to ask is how many of these units should we survey? Here again the answer revolves around our best guess as to the underlying site distribution. In most cases sites will be relatively rare phenomena, but when we do find them there will often be several close together. This, of course, is another way of saying that sites are rare and clustered events (Nance 1983; Rogge and Lincoln 1984; Schiffer et al. 1978). It is not surprising, therefore, that the results of almost every survey show that the distributions of sites per quadrat (of whatever size) are positively skewed.

The results of a BLM sample survey in the San Rafael Swell of east-central Utah can be used to illustrate a more or less typical situation (Tipps 1984). Figure 6.2 presents a series of histograms of site density for three noncontiguous survey areas. All show positive skewing. Many quadrats contain no sites and progressively fewer quadrats contain larger numbers of sites.

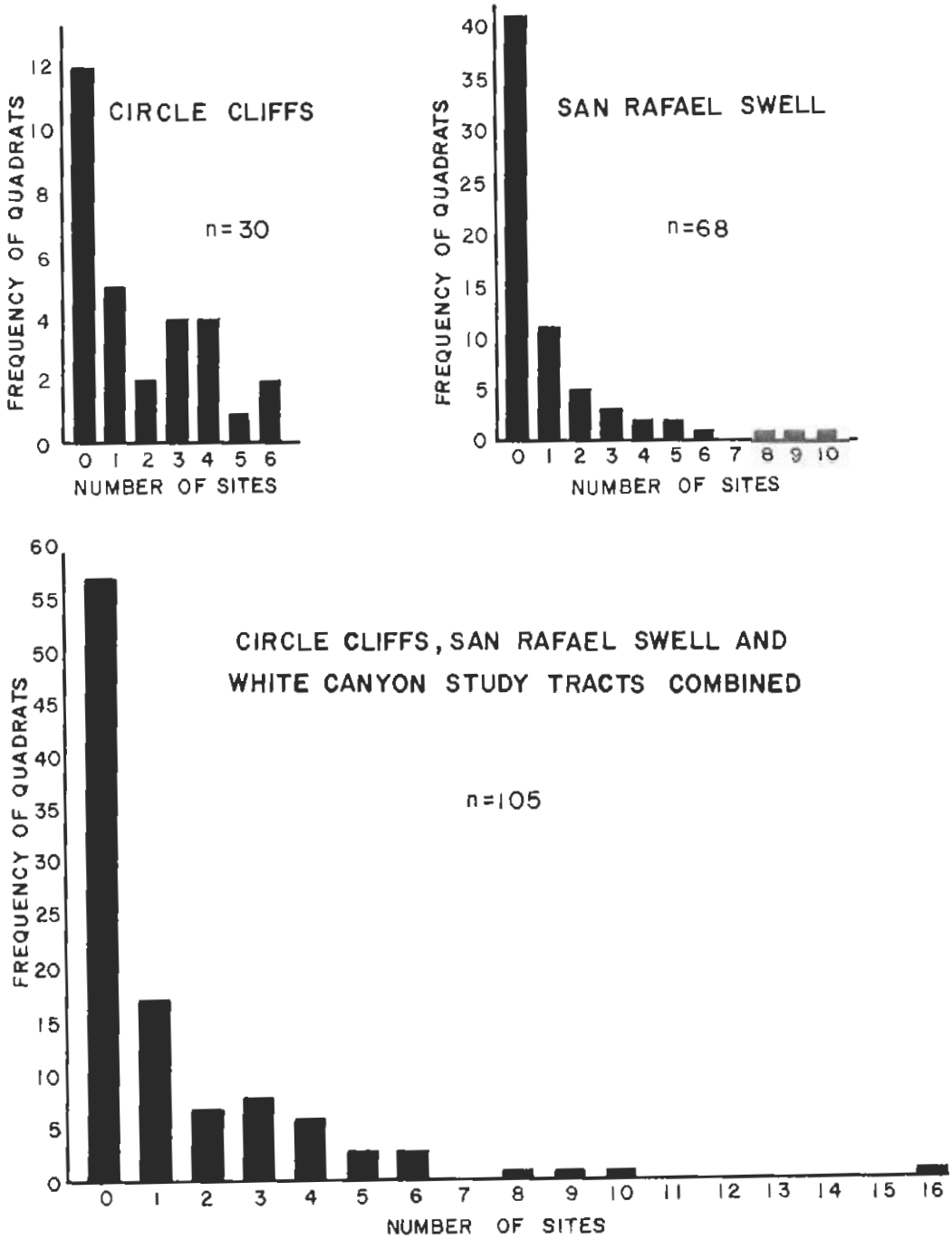


Figure 6.2. Bar graph of the frequency of quadrats by the number of sites, showing the positively skewed distribution (after Tipps 1983:Fig. 7.1).

Let us assume that the distribution of sites per quadrat for each sampling population as a whole is also positively skewed. The Central Limit Theorem (Hayes and Winkler 1971:292) states that the distribution of sampling means for samples drawn from this population will still approximate normality; the average of the sampling means will equal the population means, and the standard deviation of the sample means will equal $\sigma/\sqrt{N}$ if these are based on repeated random samples of sufficiently large size. Cochran (1977:42) suggests as a rule of thumb that sample size should be greater than $25G_1^2$, where G_1 is Fisher's measure of skewness. Table 6.1 shows that for the Utah survey results plotted in Figure 6.2, only the Circle Cliff region survey meets this criterion, with a 10 percent sampling fraction. Only by combining the San Rafael Swell survey area with the Circle Cliff survey area was the researcher able to obtain an adequate sample size for the San Rafael Swell. From an anthropological standpoint, this is a questionable practice at best.

Several archaeologists have noted that adherence to Cochran's rule will usually require a very large sample size (e.g., Thomas 1975:68–70). Nance (1983:303) has suggested another method, based on Monte Carlo simulation, in which a hypothetical population distribution is created on the basis of the sample data. Repeated sample selection from this population then allows for a thorough examination of skewness.

The task of selecting an appropriate sample size from a skewed population distribution becomes even more difficult when the subject of interest shifts from the sample unit to the site. In this case only those units that contain sites are of importance. Thus, it is not the total number of sample units but the total number of sample units *minus* the number of sample units without sites that will determine the size of the survey. The problem then is to estimate how many units will have to be surveyed before an adequate number of clusters is obtained. The work of a number of archaeologists and human geographers suggests that fitting of discrete probability distributions to supposed settlement distributions may be a useful approach to this problem (e.g., Clarke 1977; Cliff and Ord 1973; Dacey 1964; Harvey 1967; Hodder 1977; Hodder and Orton 1976; Hudson 1969; King 1969; Wood 1971). For

TABLE 6.1.

Skewness values and sample sizes by study tract

Study Tract	*Skewness Value*	*Requisite Sample Size* ($n > 25G_1^2$)	*Actual Sample Size*
Total 10 percent sample			
Circle Cliffs and San Rafael Swell	1.97	>97	98
Circle Cliffs			
10 percent sample	0.82	>17	30
San Rafael Swell			
10 percent sample	2.47	>153	68
White Canyon	2.47	>153	7

Adapted from Tipps 1984:332, Table 38

example, Wood (1971), Hodder and Orton (1976), and Nance (1983) argue that, because sites are often rare and clustered events, the pattern of site densities in many regions can be described reasonably well by the negative binomial distribution. The negative binomial is described by two parameters, the arithmetic mean and a positive exponenr (k; see also Chapter 5). For a region about to be surveyed, an archaeologist can arrive at an estimate of average site density by using the results of surveys in nearby regions with similar environments. The positive exponent, k, can be estimated in a variety of ways (see Bliss 1953). The usual approach in archaeology is first to arrive at some estimate of the sample variance and then to calculate k using the equation

$$k = x^2/(s^2 - x)$$

where x and s^2 are the mean and variance of the sample (Nance 1983; Wood 1971). Once the parameters are defined, the probabilities of obtaining a certain number of sites per unit can be calculated in a straightforward manner using the probability generating function

$$P(x) = \frac{(k + x - 1)!}{x!(k-1)!} \left(\frac{R^x}{q^k}\right) \quad \text{for } x = 0, 1, 2, \ldots$$

$$= 0 \text{ otherwise}$$

where $R = p/q = (m/k + m)$.

Nance (1983:334–335) has provided an example of how the negative binomial distribution can be used to determine how many units should be surveyed. Using sample estimates for x and s^2 from a simple random sample survey of 31 quadrats in the Upper Hat Creek region of British Columbia, Nance calculated the parameters of the negative binomial distribution, x and k. For a given quadrat size, then, he could predict the number of "empty" quadrats that would be surveyed, the number containing one site, the number containing two sites, and so on. He found that the negative binomial distribution fit the observed site distribution very closely (Nance 1983:335, Tables 8.8 and 8.9). This fit was expected since the predictions were being compared with the data from which they were derived, but the results indicate the potential of this and other probability distribution functions for indicating approximately how many empty units are likely to be found for a given sample size. By extension, if a reasonable estimate of the probability that a unit will not contain a site can be calculated, we can also determine the number of units that would have to be surveyed in order to obtain a specific number of units containing sites. For example, if the probability of any survey unit being empty is 0.50, then in order to obtain 30 units that contain sites we would need to survey

$$30 = n - (\text{number of empty units})$$

$$30 = n - (n)(\text{probability of an empty unit})$$

$$30 = n - (n)(0.50)$$

$$60 = n$$

Nance (1983:338) has pointed out that the magnitude of the empty unit problem is likely to vary widely from region to region. The problem will be even worse when interest focuses on specific site types as opposed to sites in general. At this point we have two choices. We can either adopt very large sampling fractions or try to reduce the spatial heterogeneity exhibited by most site distributions.

The primary means by which archaeologists have attempted to reduce heterogeneity in site distributions is through stratification of the sample universe. Often archaeologists divide a sample universe on the basis of criteria that they believe may have influenced site location or that they believe can serve as a proxy for such influence. Common criteria include soil type, vegetation zone, physiographic unit, or any combination of the above. In many instances the resulting areas are simply viewed as separate sample universes. For example, Thomas (1975:65) divided the Reese River region into three units on the basis of biotic communities, and the resulting subdivisions were viewed as separate sample universes. In order to draw a 10 percent sample of the entire region, Thomas actually selected 10 percent of the sample units in each sampling domain by means of a separate simple random sampling procedure.

The main advantage of this approach is that it ensures that all regions get proportionally equal coverage. Further, because simple random sampling was conducted in each region, parameter estimates can be computed for each stratum using formulas designed for such sampling. If interest focuses on estimates for the entire sampling universe (i.e., the areas encompassed by all strata combined), however, then computing these estimates is somewhat more involved. For example, to estimate the standard error of the sample mean derived from a simple random sample, the following formula is used:

$$SE = \frac{s}{\sqrt{N}} \sqrt{1 - n/N}$$

where SE is the standard error, s is the standard deviation of the sample, n is the sample size, and N is the size of the population. The standard error of the sample mean derived from a stratified random sample is calculated as

$$SE_{strat} = \sqrt{\frac{\sqrt{\Sigma(n_i s_i)(1 - n/N)}}{n^2}}$$

where SE is the standard error of the stratified sample, n_i is the number of cases chosen from Stratum i, s_i is the standard deviation in Stratum i, n is the total number of cases chosen, and N is the total number of cases in the population.

The standard error is clearly easier to calculate for simple random samples than for stratified random samples. The temptation is to make the assumption that the variability within and between strata is approximately the same and thus proceed with calculations as if the sample were a simple random one. The problem with this approach is that each variable being measured may be characterized by different levels of variability in the strata and different degrees of correlation with the criteria

used to create the strata (Dixon and Leach 1978:17). The net result for most variables is that the standard error, as computed by the simple random sample formula, is overestimated. This is true even if the same sampling fraction is used for each stratum.

The stratification approach described above is best suited to relatively large areas for which our information about site location is limited. Many Bureau of Land Management Class II coal lease inventories in the Rocky Mountains fit this description. These management-defined universes can cover well more than 100,000 acres and contain portions of several river basins. While at the outset archaeologists may not be in a position to define strata that covary with site distributions, they may be able to suggest that each major river basin could have encompassed a separate settlement-subsistence system. Failure to divide the region into natural units could lead to oversampling in some regions and undersampling in others and thus to rather poor parameter estimates.

An alternative to this type of stratification is systematic sampling. In the latter design, survey units are selected at set intervals, with the first unit usually being chosen by a random process. Several experiments with archaeological data have shown that systematic sampling can lead to relatively precise parameter estimates (Judge et al. 1975; S. Plog 1976; Sanders et al. 1979). The main disadvantage of systematic sampling is that the approach is liable to miss patterns in the underlying distribution that exhibit periodicity. Statistically, a systematic sampling design is somewhat more difficult to evaluate than a random design because bias can only be estimated (Cochran 1977; Read 1975).

Discussions of sample stratification usually do not refer to definition of separate universes. Generally, stratification means subdividing a sample universe into two or more strata and then selecting different proportions of each stratum for observation. When the population exhibits uneven spatial variability, as in the case of clustered elements, such as sites, an areal stratification scheme that samples the strata in proportion to their estimated variance will, if done correctly, lead to more precise parameter estimates than simple random sampling, systematic sampling, or stratified sampling with proportional allocation (Cochran 1977:99–103). Let us assume, for example, that a region consists of two vegetation zones, 100 km² of piñon-juniper forest and 100 km² of sagebrush. Further, let the population value for site density in the piñon-juniper zone be four sites per square kilometer with a variance of three, and the site density in the sagebrush zone be two sites per square kilometer with a variance of 0.75. A 10 percent sample of the 200 km² region using 1 km² survey units would result in the survey of 20 units. Under a simple random sampling approach, each unit selected has a 50–50 chance of being located in the piñon-juniper forest and a 50–50 chance of being in the sagebrush zone. Using a binomial distribution we can calculate the probability of selecting a specified number of sample units in one of these zones as

$$P(r) = \binom{N}{r} p^r q^{N-r}$$

where $P(r)$ is the probability of selecting r survey units, N is the total number of survey units selected, p is the probability of selecting a survey unit in the zone in question, and q equals $1-p$.

Table 6.2 lists the probabilities of selecting exactly 0, 1, 2, . . . 20 units in one of the vegetation zones. The most likely outcome is that of obtaining 10 survey units in each zone, which will occur approximately 17 percent of the time. The chances of obtaining distributions of 9–11, 8–12, or 7–13 are relatively good, with the 7–13 distribution occurring about 15 percent of the time. While it is true that over many samples a relatively even split can be expected, for any one sample there is a fairly good chance that one zone will be overtepresented and the other underrepresented. Given the population values, a simple random sample will lead to rather imprecise estimates. That is, sample estimates of the population values are likely to fluctuate very widely and thus to be associated with large standard errors.

Sampling each zone proportionally will not greatly affect this situation. In our example, if we were to treat each zone equally, exactly 10 units in each would be surveyed. For the sagebrush zone this might be sufficient, but given the large variance in the piñon-juniper zone such an approach would still lead to rather imprecise estimates. In this situation what we really want to do is to survey more units in the piñon-juniper zone than in the sagebrush zone. How many more? That

TABLE 6.2.

Probability of selecting a specified number of survey units in particular zone under simple random sampling, $p = 0.50$

Number of Survey Units Selected	*Probability*
0	0.0000009
1	0.00002
2	0.0002
3	0.001
4	0.005
5	0.015
6	0.037
7	0.074
8	0.120
9	0.160
10	0.176
11	0.160
12	0.120
13	0.074
14	0.037
15	0.015
16	0.005
17	0.001
18	0.0002
19	0.00002
20	0.0000009

depends on the variance of the sample mean and the cost of taking the sample. Cochran (1977:96) defines one function for computing cost as

$$\text{cost} = C = c_0 + \Sigma\, c_h n_h$$

where c_h is the cost per unit in Stratum h, n_h is the number of units observed in Stratum h, and c_0 represents an overhead cost. In archaeology, costs per unit would include such items as recording time and travel time (often the latter is represented mathematically as

$$\Sigma\, t_h \sqrt{n_h}$$

where t_h is the travel cost per unit). The objective, then, is to minimize cost for a specified variance of the stratum's sample mean or to minimize the variance of the sample mean for a specified cost.

If one is not in a position to estimate cost and is willing to assume that cost per unit is the same in all strata, then determining optimum allocation reduees to the equation

$$n_h = n \frac{N_h S_h}{\Sigma N_h S_h}$$

where n_h equals the number of cases to be selected in Stratum h, n refers to the total sample size, N_h equals the number of potential cases in Stratum h, and S_h is the variance of the sample mean in Stratum h (Cochran 1977:98). This allocation is often referred to as *Neyman's allocation* (Neyman 1934). For our previous example, using Neyman's allocation we would obtain the following results:

$$n_{\text{piñon-juniper}} = \frac{20(100 \times 3)}{(100 \times 3) + (100 \times 0.75)} = 16$$

$$n_{\text{desert scrub}} = \frac{20(100 \times 0.75)}{(100 \times 3) + (100 \times 0.75)} = 4$$

The point here is that if sample size and sample fraction are relatively small, then use of prior knowledge about the nature of the phenomenon to be modeled may be the best way to obtain the precise estimates needed for modeling. The use of such information in archaeological modeling has been rather limited, perhaps because many archaeologists believe that they are not in a position to offer even good guesses as to the underlying population values.

One approach to circumvent this problem is to perform a pilot study. Fot instance, if we were conducting a 10 percent sample survey of a national forest for the purposes of estimating site density, one strategy would be to select a 10 percent simple random sample of predefined areal grid units. This approach, while perhaps meeting the assumptions of sampling theory, many times leads to very poor results.

Estimates are often not very precise, and one is left with the feeling that for all the rigor we have still not learned very much. A better approach might be to assume that site locations covary with certain mappable features (e.g., soils or landforms). This assumption could be tested by some type of probing or purposive survey (see next section) and/or by a relatively small, simple random sample survey. Based on the results of this survey, specific criteria could be defined that would lead to a useful stratification of the region.

If a small simple random sample survey had been conducted, then the sample could be *poststratified*; that is, each of the units surveyed during the pilot simple random sample could be reclassified into one of the newly defined strata. Cochran (1977:134) notes that poststratification is almost as precise as proportional stratified sampling in providing parameter estimates as long as the samples in each stratum are reasonably large (say, more than 20) and the effects of errors in the stratum weights can be ignored. Basically, care must be taken to ensure that the final sample matches the population in important respects. If, for example, access to survey was denied on private land along river bottoms, the sampling frame might have excluded a high proportion of a certain site type. Simply giving added weight to the sites of that type that *were* included in the survey may not improve the sample's estimate for the density of that site type; indeed, it may make it worse (Dixon and Leach 1978:21).

If we can justify poststratifying the sampling universe, then we can use the variance estimates for each stratum to determine the optimal allocation of cases for the second stratified random sample. Chances are extremely good that even though the parameter estimates of the stratified random sample would be based on a smaller number of cases (that is, assuming that the pilot study and the stratified random sample survey together covered 10 percent of the region) the gains made by stratifying the region would still lead to more precise estimates than those based on a single 10 percent simple random sample.

As a final note, we want to point out one more serious problem with using data derived from stratified random sampling to develop predictive models. Generally, when a multivariate pattern-recognition model is developed some type of commercial software is used. All statistical software packages of which we are aware assume simple random samples—that is, the variance-covariance matrices are computed as if the data were obtained through a simple random sampling procedure. If stratified random sampling was used instead, then the matrices will be computed incorrectly. The statistical ramifications of this error are not well understood, although it is clear that the variances will be overestimated. Perhaps the best approach to this problem is to write a simple program to compute the matrices correctly and then use these matrices as input to the desired algorithm.

Purposive Selection

One of the main objectives of collecting new data for predictive modeling is to make certain that no magnet site is missed. Many such sites will have been recorded

prior to the survey or will at least be known to local informants. In this situation all that need be done is to verify the site's location and the nature of its surface assemblage and environmental context. In cases where there is reason to believe that not all magnet sites are known, survey strategies that maximize the chances of finding sites in this category must be designed. There are two options. When there is evidence that important centers were distributed according to some predictable feature of the natural landscape, such as at a regular interval along a major river or at the confluence of major watercourses, specific areas can be picked to survey. A second approach, which is especially useful in regions occupied by complex societies, is to use some type of remote sensing information. Because regional centers tend to be the largest and most complex sites in a given area, they can often be detected on aerial photographs (see Chapter 9). Another technique that enables the archaeologist to cover extensive ground areas in short periods of time is aerial survey from a small-engine aircraft or a helicopter. In this regard, Rogge and Lincoln's comments concerning the Tucson Aqueduct surveys (described in Chapter 3) are particularly appropriate.

> Our Tucson Aqueduct case indicates that we did learn a great deal with each new survey but implies that our predictive models were not particularly robust. Neither did conducting the surveys exactly "by the book" ensure meaningful input into our planning process. . . . If we were to start the Tucson Aqueduct series of surveys over today with the 20–20 vision of hindsight, we might decide to spend a few days with a helicopter looking for platform mounds and do nothing more until a route was selected [1984:19].

The two approaches are not mutually exclusive. Indeed, in one of the most intensive archaeological surface surveys ever conducted, Millon (1972:11–12) had the entire confines of the city of Teotihuacan photogrammetrically mapped to reveal low-lying mounds, which are often the remains of architectural features. The maps were then used to guide subsequent fieldwork.

Research designs can incorporate both purposive selection and probability sampling. During the Tucson Aqueduct surveys, for example, had the Bureau of Reclamation conducted a helicopter survey and found the three platform mounds, a stratified random sample survey could have been conducted. Three sampling strata consisting of arbitrary 10 by 10 km grids centered over each platform mound and a fourth stratum representing the remainder of the survey universe could have been defined, with the surveyors covering relatively high sampling fractions in each Hohokam community stratum and a much lower fraction of the remaining region.

Finally, it is important to point out that even with the best of sample survey designs, magnet sites will still be missed. Some have argued that this is exactly why sample surveys and predictive models should not be used. In an absolute sense, these critics are right; present models make more mistakes (especially gross errors) than anyone is willing to accept. But blind 100-percent surveys are not necessarily the answer. Complete inventory surveys that have no theoretical foundation often end up adding little to our understanding of prehistory. Further, depending on the

field methods (i.e., crew spacing, recording technique, etc.), "100-percent" surveys can easily miss all types of sites, including magnet sites. In short, regardless of whether the sampling fraction is 10 percent or 100 percent, there is no substitute for a well-thought-out survey design that is grounded in some theoretical foundation.

Depositional and Postdepositional Processes

The final class of data needed for the creation of a predictive model concerns the processes affecting site detection and site survivability. From a research perspective, it is important to be able to predict areas where sites probably were located but where evidence of past activities has been destroyed by natural processes. While negative evidence may not be very helpful in substantiating hypotheses about settlement location, proper geomorphic interpretation may be critical if we are to avoid incorrect rejection of a hypothesis because of the lack of cultural remains. From a management perspective, it may be less critical to model site destruction that has resulted from natural processes, but it is still necessary to model locations of sites that are intact but not visible on the surface. Buried sites are perhaps the land manager's worst nightmare. Often they are not found in the course of usual cultural resource studies and are only detected after construction or development has begun. The mitigation of adverse effects on buried sites often ends up costing much more than the expenses of the archaeology alone.

To find buried sites the first step is to detect and trace paleo land surfaces suitable for habitation. This task properly falls into the field of geomorphology. While archaeologists have worked with geomorphologists for many years (e.g., Butzer 1971, 1982; Davidson and Shackley 1976; Hassan 1979; Haynes 1968; Haynes and Agogino 1966; Jacobsen and Adams 1958; Martin and Klein 1984; Saucier 1974), this working relationship by and large has not been transferred into the area of predictive modeling. Geomorphic fieldwork should ideally precede at least one stage of archaeological fieldwork. The results of the geomorphic analyses are often presented as maps of paleo land surfaces that specify areas where buried sites are likely to be found. If such studies were carried out in conjunction with archaeological surveys, areas designated by the geomorphologist could be examined with subsurface tests.

The issue of subsurface testing on surveys has recently received considerable attention (e.g., Krakker et al. 1983; Lightfoot 1986; McManamon 1984; Nance and Ball 1986; Wobst 1983). Most of this interest stems from research in forested areas where the ground surface is obscured. In these situations visual inspection of the surface greatly underestimates the numbers of sites and leads to highly skewed locational patterns. If these factors are not taken into account, then statements about settlement patterns that implicitly assume that the observed sites are representative of site locations in general are likely to be highly inaccurate.

Thus far most research on discovering buried sites has focused on sites that lie on or near the surface. The approach that has gained widespread acceptance in this

situation is to space small subsurface probes, usually in the form of shovel pits or test pits, at a set interval along a survey transect. Nance and Ball (1986) have shown that the likelihood of discovering sites with subsurface tests varies directly with the artifact density and the size of the sites. Another key variable in determining site discovery potential is the intensity with which the fill of the test is inspected. The probability of site discovery increases dramatically with a shift from visual inspection of the fill to screening of the fill, and the probability increases still further as the size of the screen mesh decreases. It is worth pointing out, however, that even with a small interval between subsurface tests and screening through fine mesh, the likelihood of missing small, low-density sites is usually very high.

The problem of buried sites is not confined to forested areas or regions where the ground surface is obscured. Geomorphic changes can lead to buried sites in areas with good surface visibility. For example, in some desert areas of the American Southwest, remains of the Hohokam culture (ca. AD 200–1450) can be found on the surface. Pedestrian surface survey results usually correlate fairly well with intact subsurface deposits of this age. Remains of the preceding Archaic and Paleoindian periods, however, are not generally found on the surface. Sites associated with these periods tend to be found in deep erosional cuts or as the result of modern land disturbance or construction. Thus, an interpretation of negative results of pedestrian surveys in these regions as meaning that no Archaic or Paleoindian sites lie in the survey area involves an inaccurate and unjustifiable logical leap from the surface to the subsurface.

The problem of buried sites is fairly widespread and will always have to be taken into account when designing surveys to build predictive models. One approach is to use the results of a geomorphic analysis as a means of stratifying the area. The paleo land surfaces identified could each be assigned a relative probability of site discovery. This probability could be based on previous research, the types of depositional environments represented, or a combination of these factors. Each stratum could then be divided into grid units and a number of grid units selected for survey through a random process. Optimum allocation of the number of grid units selected in each stratum could be based on the relative probabilities previously defined. Each grid unit could then be subdivided into smaller units, with a set number of these units being selected for subsurface tests through either a random or a systematic process.

The sampling scheme described above is referred to by statisticians as *two-stage* sampling or *subsampling* (Cochran 1977). Although the statistics can become rather involved, this type of survey design can lead to unbiased and precise parameter estimates. Parenthetically, if the design is extended to sampling the subunits themselves, then it is referred to as *three-stage* sampling or *multistage* sampling. The latter term is often misused by archaeologists to refer to sampling designs that are carried out in sequential steps (e.g., conduct a 1 percent simple random sample survey of a region [step 1], stratify the region [step 2], conduct a 10 percent stratified random sample survey of the region [step 3], and so on). Although the term *multistage* seems entrenched in the archaeological literature, to avoid confusion

with other uses of the term we refer to this type of sampling design as *multistep* throughout this volume.

DATA COLLECTION IN CRM CONTEXTS

The preceding discussion of survey strategies was presented as though we already knew where we were going to survey, how we were going to survey, and how we were going to record data. In practice, these are three of the most important factors that are involved in designing a survey. In an ideal setting, all three are determined on the basis of research objectives. But surveys conducted in cultural resource management contexts are subject to a set of unique constraints that often greatly restrict the ways that these three factors can be integrated into the overall survey design. In the following discussion we examine the ways in which management needs have shaped survey design and evaluate the common responses to these needs in terms of their usefulness for model building. These issues will be discussed under three specific topics: survey universe, survey intensity, and data recording.

Survey Universe

Ideally, the selection of an area to survey or from which sample units are to be selected is based on theoretical propositions underlying the research design or topics. In theory, researchers want to select a survey universe that conforms to a cultural unit. In practice, however, at best we can only approximate this situation. Cultural systems rarely have sharp boundaries. Defining where one system ends and another begins is usually impossible for ethnographers, to say nothing of the problem faced by archaeologists. Further, cultural systems change through time in nature and in size. Thus, a survey universe suitable for studying one culture may be too large or too small for examining its predecessors and its successors.

A common solution to this dilemma is to select a region that conforms to a natural unit, such as a drainage basin or an island, with the size and type of the natural unit selected depending on the research topic. At one extreme, Sanders chose the entire Basin of Mexico as the survey universe for a study of the origin of state-level societies in highland Mexico. The ensuing project lasted 15 years and involved about 50 field months of actual survey (Sanders et al. 1979:19). Most projects are not nearly as large as the Basin of Mexico survey, but in all cases the selection of a survey universe is a compromise between two opposing criteria. On the one hand, we want a region that is large enough so that we can reasonably argue either that the remains of the prehistoric settlement systems that characterized the area are contained within the survey universe or that all major components of those systems are at least represented. On the other hand, we want the survey universe to be as small as possible, thereby allowing us to maximize our survey effort.

For surveys conducted in a cultural resource management context, the survey universe is most often defined not by archaeologists but by land managers, who must take into account many factors that have little to do with archaeological research. In some cases the survey universe will encompass one or more natural units, but usually it will not. It is appropriate, therefore, to consider the implications that management-defined survey universes have on building predictive models.

To illustrate these issues we will use the example of a Bureau of Land Management cultural resource management project in the San Rafael Swell region of east-central Utah that was mentioned earlier. The San Rafael Swell is an elongated anticline approximately 110 km (50 mi) long and 50 km (23 mi) wide. Since 1979 the Bureau of Land Management has sponsored six major survey projects in the region (Table 6.3). All of these projects were designed as probabilistic sample surveys of management-defined survey universes encompassing different portions of the swell. Of the more than 550,000 ha (ca. 1,360,000 acres) comprised by the San Rafael Swell, more than 10,000 ha (ca. 25,000 acres; 1.8 percent of the total area) were inventoried as part of these seven projects.

The modeling efforts carried out in conjunction with these survey projects mirror the general trends in predictive modeling. The first locational analyses consisted of univariate and bivariate correlations between site location and specific

TABLE 6.3.

Estimated density of prehistoric sites in project areas in and near the San Rafael Swell

Project Area	*Quadrat Size (acres)*	*Number of Survey Quadrats*	*Number of Prehistoric Sites*	*Mean Number of Sites per Quadrat*	*95% Confidence Interval*
San Rafael Swell 10 percent sample	160	68	80	1.18	±0.50
Central Coal Project Muddy Planning Unit (Hauck 1979a)	160	22	50	2.27	n/a
Central Coal Project Summerville Planning Unit (Hauck 1979a)	160	15	19	1.27	n/a
Southern Coal Project Huntington Planning Unit (Hauck 1979b)	160	10	7	0.70	n/a
Central Coal II, Tract II (Thomas et al. 1981)	160	31	7	0.22	±0.17
Central Coal II, Area III (Thomas et al. 1981)	160	11	101	9.18	±3.92

Adapted from Tipps 1984:139, Table 42

environmental variables (Hauck 1979a, 1979b). Significant associations between site locations and environment were then combined into overlay models (Thomas et al. 1981). Finally, most recent attempts use sophisticated multivariate discriminant function analysis and hierarchical clustering models (Tipps 1984).

None of these models has been a very good predictor. This is not a result of glaring errors in the derivation of the samples or the application of the statistics. Instead, the poor accuracy rate appears to result from "tunnel vision." Each model was based on an inductive, pattern-recognition approach that viewed the survey universe as the only region of interest. Even a casual glance at Table 6.3, however, indicates that wide fluctuations in mean site density exist. This is probably also true of the sample variances, but except for a few cases, these are not published. These variations are probably caused by settlement and subsistence practices that are regional in scope. Thomas and his colleagues appear to recognize this situation and state that

> The Central Coal II Class II Inventory was designed as a 10 percent simple random sample of three sampling universes (Study Tracts I, II, and Area III). The notion underlying this type of approach is that the survey results of the sampled portion of each tract can be generalized over the entire tract. This method may be useful for evaluating site sensitivity in Tracts I and II. But . . . Area III does not appear to be a self-contained cultural unit. Settlement in this area seems to be directly related to practices in the adjoining regions. . . . Trying to generalize the results of the sampled portion of Area III to the entire tract is apt to be misleading. Many of the critical features in the settlement system clearly were not included in the sampling universe. Instead of developing a statistically valid model that makes little logical sense, it seems far preferable to create an internally consistent model of the settlement system that can then be used to evaluate the Area III position of the system and thus to predict areas of site sensitivity [Thomas et al. 1981:199].

Thus, while a 10 percent simple random sample of Central Coal II, Area III might yield a representative sample of *spatial units* for that area, it is quite possible that regional patterns in the settlement system would go undetected on the basis of this sample. Even though the parameter estimates might be reliable, predictive models based only on patterns discernible within the sample universe are, in this case, likely to yield disappointing results.

There is no easy solution to this problem. To build useful predictive models the researcher must have reason to believe that the survey universe conforms to a cultural unit, or failing this, he or she must use a defendable proxy, such as a natural unit. If we must use management-defined survey universes, then it is critical that the fit between an appropriate cultural or natural unit and the arbitrary universe be assessed. In addition, the resulting model must take into account the position of the resources in the survey area relative to the larger settlement-subsistence system, and it must incorporate regional factors affecting settlement location.

Designing a research strategy to accomplish this task may involve some restructuring of many cultural resource management programs. To use the BLM as an example, one solution would be to subdivide each district into natural units.

Instead of building a model for each major coal lease project, archaeologists might build a predictive model for each individual natural unit within the district, with the model being periodically refined as more data become available. Thus, the data from a large coal lease project might be better used in several regional predictive models instead of one ad hoc, project-specific model. Such a program would require that each project carried out within the district emphasize the collection of comparable data. While it is possible to combine probabilistic samples, this requires considerable statistical expertise. A much more serious problem is that of ensuring that the entire sampling universe is adequately covered. The usual government policy is not to survey privately owned land. In many areas private property covers much of the "desired" land, such as the bottomlands in a river valley or the elevated, well-drained soils in a deltaic plain. Many sites, including a high proportion of magnet sites, will often be found on private land. If we eliminate such areas from our sampling universe, our ability to predict site location will be greatly hindered.

Certainly there are many problems involved in designing cultural resource management projects that focus on culturally meaningful study areas, but projects that emphasize development of ad hoc models for arbitrary units are clearly as responsible for the poor showing of these models as anything else. Unless this focus changes so that the models can be related to cultural phenomena, it is unlikely that the results will improve.

Survey Intensity

Without doubt the single most important factor affecting the number of sites located on a survey is the effort made to find them. Survey intensity can be measured in terms of the ratio of person-days to square miles surveyed or on the basis of the spacing between surveyors (Judge 1981; S. Plog et al. 1978; Schiffer and Wells 1982). Regardless of the measure used, all studies to date confirm Judge's (1981:128) statement that "the more time spent in the field looking for sites, the more sites will be found."

S. Plog et al. (1978:391–393) examined the relationship between survey intensity (as measured by person-days per square mile surveyed) and estimates of site density using the results of 12 surveys conducted in the southwestern United States. They found a strong positive linear correlation between the two variables, which is to say that as survey intensity increased so did site density. Part of this relationship is a result of the time spent in recording sites and making collections once the sites are found. Thus, we would expect that as more sites are found more time must be spent in the field. S. Plog et al. (1978:393) argue convincingly, however, that this is not the whole story, that indeed, if one controls for extra time spent recording sites, a strong positive relationship still exists between survey intensity and site density. In theory, a point of diminishing returns should be reached beyond which increases in intensity do not result in proportional increases

in site density estimates; for the 12 surveys studied, however, no evidence was found that indicated that such a point had been reached (S. Plog et al. 1978:393; see also discussion in Chapter 4 of this volume).

Selection of an appropriate level of survey intensity requires careful consideration of several factors. The major consideration affecting survey intensity should be the research objectives. Is it necessary to locate "all" resources, or are we primarily interested in specific types of sites? For example, the Basin of Mexico survey project discussed earlier was designed to recover "a variety of data on where people had lived during the pre-Hispanic past in the survey area" (Sanders et al. 1979:15). The surveys, therefore, focused on habitation sites and made no attempt to identify more ephemeral, limited-activity loci. The selection of a flexible survey interval of between 15 and 17 m (Sanders et al. 1979:24) was appropriate to these objectives.

In cultural resource management contexts, surveys are rarely focused on a particular type of site, and even surveys designed to acquire data for a specific set of research objectives are uncommon. Usually the stated goal is to find "all" the resources. Such a hubric ideal can never be achieved, however, and what is really meant by "all" is some very high proportion of the recoverable resources.

Selection of an adequate survey intensity also depends on the nature of the resource base and the prevailing natural conditions. As discussed in Chapter 4, the latter directly influence our ability to detect archaeological material, a factor categorized by Schiffer and others as *visibility* (Schiffer and Gumerman 1977:186–187; Schiffer and Wells 1982:349; Schiffer et al. 1978:6). In general, high visibility means that if cultural remains exist on the surface an observer should be able to see them. High-visibility areas generally have sparse vegetation, e.g., deserts, beaches, or plowed fields. Low-visibility areas have masked or obscured surfaces. Pedestrian surface survey techniques yield poor results in these areas and must be supplemented by subsurface investigations, such as shovel tests or test pits, or by techniques that expose the surface, like raking or plowing.

Cultural factors affecting the likelihood of site detection include site size, site obtrusiveness, site distribution, and surface artifact density. In general, larger sites have a better chance of being found than smaller ones; sites with high surface-artifact densities are more likely to be seen than those with sparse or no surface expression; and sites with obtrusive features, such as mounds or masonry, are easier to find than sites lacking such features. While these generalizations may seem to be self-evident, they have important implications for the model-building process. Previously it was argued that in order to construct a successful predictive model we need *(a)* to have reliable estimates of a number of parameters associated with site location, *(b)* to locate all or most of the magnet sites, and *(c)* to assess the effects of depositional and postdepositional processes on site visibility. To discover magnet sites, large areas must be covered, but often these areas can be surveyed at very low intensities without affecting the result. In contrast, accurate parameter estimation for less-obtrusive sites requires a much higher level of effort per area surveyed. Given these competing requirements, several archaeologists have recently advo-

cated multistep survey designs in which different types of data are acquired at different stages (e.g., Doelle 1976, 1977; Schiffer and Wells 1982; Schiffer et al. 1978).

Since survey intensity directly affects the rate of site discovery, one would think that this issue would weigh heavily in the evaluation of proposed survey strategies. In practice, this is rarely the case. Many scopes of work specify the interval between surveyors; the types of subsurface tests, if any, that will be conducted; and the information that is to be recorded on each site. The rationale for providing these fixed specifications appears to be that this will ensure that all contractors bid on the same work. While the objective is understandable, it is important that the land-managing agencies realize the effects of this decision on the model-building process. When these aspects of the survey methodology are prespecified, survey intensity becomes a parameter rather than a variable. Thus, what is probably the single most important factor affecting the power of any predictive model is being arbitrarily set by the managing agencies for reasons that have little to do with archaeology.

The point is that selection of the survey intensity is a critical and integral step in the model-building process. The choice should be based on fieldwork and subject to testing and refinement, as well as to changes when the research objectives change. One contribution that the managing agencies can make to the accuracy rate of predictive models is to allow survey intensity to be set on the basis of archaeological considerations rather than procurement procedures.

Data Recording

In the preceding section we discussed some of the factors affecting the number and types of sites discovered. Yet we side-stepped perhaps the most important issue—what is a site? To a large extent, site definition is actually an issue of data recording. That is, we need to define consistent and replicable criteria by which space can be partitioned into those areas that we want to call sites and those that we do not. Traditionally, this issue has not been problematic. Archaeologists tended to focus on large sites with discrete boundaries, such as masonry pueblos or earthen mounds. In the last decade, however, some researchers have focused on loci where evidence of cultural activity is more ephemeral, such as isolated finds or low-density artifact scatters, and it has become clear that these phenomena can be quite important to our understanding of the prehistory of a region (e.g., Doelle 1976, 1977; Goodyear 1975; Teague and Crown 1983; Thomas 1975).

This awareness of the continuous aspects of the archaeological record has led a number of archaeologists to question the utility of the site concept (Dunnell and Dancey 1983; Ebert et al. 1984; Thomas 1975; see also the discussion by Ebert and Kohler in Chapter 4). These investigators have rightly pointed out that sites do not behave; rather, people behave, and these behaviors have a spatial dimension that in no way correlates with discrete boundaries on a one-to-one basis. The problem of site definition is directly analogous to the "community boundary" issue, which has

been extensively debated in social anthropology for the past 50 years (Bell and Newby 1974; Galeski 1972; Goodenough 1966; Leach 1961). The central point of this issue is that anthropologists have taken the village or settlement as their unit of analysis even though they recognize that people living in a village may work outside the village, may own land outside the village, may travel outside the village, and may have relationships with people living in other villages. The basic question, then, is at what point does the researcher set boundaries for the analysis and treat the resulting unit as an object of scientific inquiry? While no absolute answer has emerged, most anthropologists have used a spatial aggregate (whether it be a village, town, or city block) as the unit of analysis. They argue (at least implicitly) that the people within this unit are more similar to each other than they are to people living outside the unit and/or that they have more relationships with each other than they do with outsiders. In archaeology, Chang (1967, 1968) has put forward similar arguments in favor of using the settlement, defined as a single component, as the unit of analysis. As many critics of Chang's approach have pointed out (Binford 1968; Clarke 1968:648), however, components can only be defined after the assemblage has been analyzed.

While it may be obvious at the time of a survey that a mile-long lithic scatter represents multiple occupations, one still has to deal with the problem of recording it. Should an attempt be made to define discrete loci as separate sites, or should the entire area be labeled one site? Further, if the artifact scatter extends beyond the survey unit, should the entire scatter be recorded or only the portion within the unit? On the last point most archaeologists would agree that if part of a site is located in a sample unit, the entire site should be recorded. In practice, however, there are instances, such as coastal shell middens and lithic quarry sites, that can easily extend into two or more sample units and for which any boundary is somewhat arbitrary.

There is no easy way to answer these questions in the abstract. Many agencies and institutions have tried to resolve them by adopting arbitrary criteria, such as a minimum of five flakes per 5 m^2, for site definition. This practice is not without its problems, and it has important implications for model building. For example, consider two areas, one in which five flakes were found in a 5 m^2 area and another in which four flakes were discovered in an area of the same size. Under the arbitrary definition given above, the first area would be recorded as a site and the second as containing four isolated finds. During the development of a predictive model, isolated finds are usually either ignored or given the same weight as sites. For the example above, this would result in a model that would either incorporate five sites, four of which are in exactly the same environment, or one site, with the area containing the four isolated finds being considered a nonsite. Does this make sense in terms of human behavior? Most likely it does not.

The decisions as to what will be designated as a site and how that phenomenon will be recorded must therefore be based on the issues being addressed. In the case of the Basin of Mexico survey discussed above, interest focused on the development of complex societies, and the survey crews concentrated on finding habitation sites

(Sanders et al. 1979). In contrast, in the Reese River survey Thomas (1975) was interested in settlement and subsistence patterns of Great Basin hunters-and-gatherers, and the basic unit of analysis shifted from the site to the artifact.

While the definition of a site, or more precisely of the unit of analysis, must necessarily be related to the research question being addressed, it is also critical that resources be recorded in a replicable and consistent fashion. Ideally, we should be able to record resources in a way that is independent of how a "site" is defined. For many state and federal agencies the site itself is little more than a bookkeeping device for maintaining accurate records. For these purposes an arbitrary definition will suffice. The problem, then, is to find a way to fill out site records for agencies using one definition, while retaining the capability to manipulate the data according to any of a number of other definitions.

One approach to this problem is to view archaeological data as a series of hierarchically arranged dimensions. The scale at which data are collected will determine in what ways they can be used in subsequent analyses. Data collected at more specific levels can usually be aggregated to express information at a higher level, but the reverse is not true. For example, data on artifacts can be grouped to provide characteristics of features or sites (such as counts of different artifact types), but information collected at the site level cannot be used to derive information about artifacts or features found within sites.

In view of the ongoing debate about the desirability of conducting "siteless" archaeology (Dunnell and Dancey 1983; Ebert et al. 1984; Chapter 4 of this volume) within the context of predictive modeling, it may be worthwhile to explore the possibility of collecting field data in several hierarchical levels, with the data being organized in such a way that relationships between levels are easily recoverable. That is, data could be collected at the levels of *(a)* the survey units, *(b)* the sites (however one might choose to define them), *(c)* the different activity areas or features within sites, and *(d)* individual artifacts, whether from particular features or as isolated entities. Identification of the survey unit in which sites are found, the site in which features occur, and the features with which artifacts are associated (use of *pointers* to different levels in the hierarchy) would permit data from more specific levels to be aggregated or combined in order to provide variables containing information about the next higher level in the hierarchy. Durand and Davis (1985) have recently reported a similar scheme, which they designed to manage archaeological resources in Nevada. Other states, such as Hawaii, also have similar data base systems.

Table 6.4 presents a hypothetical example of this approach. It depicts a four-level hierarchical design extending from the survey unit (the highest level) to the artifact (the lowest level). In Table 6.4 the relationships between different levels in the hierarchy are maintained by labels on each successive record that identify, in turn, the survey unit, the site, the feature, and the artifacts (which are either isolated finds or parts of features). A particular level in the hierarchy can take on a null value in order to accommodate features that are not associated with sites in the traditional sense, as well as isolated artifacts. In practice, the number of hierarchical

TABLE 6.4.

Hypothetical four-level hierarchical field data file

Survey Unit 1				(survey unit data)
1	Site 1			(site data)
1	1	Feature 1		(feature data)
1	1	1	Artifact 1	(artifact data)
1	1	1	Artifact 2	
1	1	1	Artifact 3	
1	1	Feature 2		
1	1	2	Artifact 1	
1	1	2	Artifact 2	
1	1	2	Artifact 3	
1	Site 2			
1	2	Feature 1		
1	2	1	Artifact 1	
1	2	1	Artifact 2	
1	Site 3			
1	3	Feature 1		
1	3	Feature 2		
1	3	2	Artifact 1	
1	3	2	Artifact 2	
1	0	Isolated Feature 1		
1	0	Isolated Feature 2		
1	0	2	Artifact 1	
1	0	2	Artifact 2	
1	0	0	Isolated Artifact 1	
1	0	0	Isolated Artifact 2	

levels, their labels, and other details of implementation would be the responsibility of the system designer/user.

This methodology would make it possible to deal with cultural remains occurring either in packets termed *sites* or as individual items varying in density across the landscape. In the latter instance, if the spatial coordinates of artifacts have been recorded, their locations could be entered into density-contouring algorithms or *k*-means analysis (Kintigh and Ammerman 1982) as the basis for activity area, feature, or site definitions. Furthermore, characteristics of lower-order records could be used in any number of ways to construct variables descriptive of higher-order entities. Artifact variables could be transformed, for example, to produce new information to describe the features or sites in which they were found. Similarly, data from features might be aggregated to characterize sites further, and data from sites might produce additional information on the survey unit in which they were located. Counts of different types of artifacts recovered from features might be

transformed to construct a new variable of artifact density for each feature, or different types of features found on sites might be tallied to build a new variable to characterize sites. It is easy to envision many other kinds of aggregated variables that could be developed from lower-level constructs to characterize higher-level entities.

A second major issue in cultural resource management that directly affects how resources are recorded is the question of whether or not artifacts should be collected from the surface. Over the past decade a "no collection" policy has become standard for more and more federal agencies. The basic reasoning behind this policy is that more information is lost by uncontrolled surface collection than is gained by having access to cultural materials in the laboratory (S. Plog et al. 1978; Schiffer and Gumerman 1977). Certainly many cultural resource inventory surveys are conducted without benefit of a research design, and in these cases collecting artifacts, especially if they will not be analyzed, serves no useful purpose. But if the development of regional predictive models, such as those advocated earlier in this chapter, were to become a major objective, then results from all surveys could be used in the process of model building. In this case, the no-collection policy would have serious ramifications.

While in theory a no-collection policy should not affect either the quality or the extent of artifact analyses, in practice there is little question that it does. In-field analysis requires a level of competence for crew members that is generally not met. When in fact the requisite expertise is assembled, the costs of in-field analysis rise to a level comparable with laboratory analysis. There is no question that, as commonly used in cultural resource management, the no-collection policy saves money. The question is, at what cost?

As is the case with so many survey decisions, the impact of the no-collection policy is different on different types of sites. For large sites with high surface-artifact densities, this policy may not have serious negative effects. Ample numbers of temporal diagnostics can usually be found on the surface without much rrouble, and even without diagnostics these sites are generally classifiable into one of only a small number of functional site types. Real problems can arise, however, when low-density artifact scatters are encountered. In such cases we usually need all the information we can get in order to even hope to define useful analytic units. Often, distinguishing criteria, such as the presence or absence of a certain chert type or the proportion of flake categories, will not have been devised at the time of fieldwork. Thus, even if crews are well trained for in-field analysis, it is simply not possible to foresee all the observations that might prove to be informative. Further, detailed, technically complex analyses, such as wear pattern or organic residue analysis, may be required to address issues of site function. These simply cannot be conducted in the field.

The no-collection policy is in part responsible for our current inability to distinguish useful site classes, and it is unlikely that this situation will change until the policy is altered. The problem of lost provenience for surface-collected materials does not necessarily call for the radical measure of prohibiting collections; contract-

ing agencies could simply require that provenience information be recorded. If a project is designed to collect data for predictive modeling, what is needed is not that we record less information but that we record more information and record it more accurately.

The issue of accuracy is central to any discussion of data recording. Basically, to develop a predictive model three categories of data are needed: *(a)* locational data, *(b)* environmental contextual data, and *(c)* cultural data. It should be pointed out that, if nonsites are to be used in the predictive model, data on the first two categories must be recorded for each nonsite location as well.

The importance of precise locational data may seem obvious for any project with the stated goal of developing a predictive model of site location. What is perhaps not so obvious is the difficulty of obtaining such data. S. Plog et al. (1978:415) cite experiments on Black Mesa in Arizona in which sites were revisited to check on locational accuracy. Considerable variation was found, with some sites located accurately and others having been plotted more than 200 m from their correct location. These problems tend to multiply as more researchers work in an area through time. In a Class I overview for the Upper Gila River in Arizona, Phillips et al. (1984) found that the same site had been recorded three separate times (twice by the same institution) and plotted in three different locations. Portions of another site had been recorded as two separate sites by survey teams who were recording only the portion of the site that fell within their project area.

Locational errors such as those cited above indicate the need for some type of error-checking program within the survey design. Ideally, such a program would include "double-blind" tests in which a second survey crew with no knowledge of the first crew's results resurveys the same quadrat. This procedure would be especially helpful for federal agencies, such as the BLM, which have placed a high priority on maintaining comparable data standards between surveys. Double-blind tests allow us to assess locational accuracy, and because two crews record the same resources, they also permit us to examine variation in the other aspects of site recording discussed below.

Another approach to assessing the accuracy of data recording is through random spot checks. Such a program would ensure that sites are recorded accurately, but it will not assess whether sites were missed. A third approach, often used on large surveys, is to use separate survey crews and recording crews. Survey crews mark encountered sites on a map (and, if possible, in the field) and the sites are then visited by the recording team. This approach has the advantage of providing a check on recorded site location and of improving the consistency of data recording. The recording crews usually have fewer people than the survey crews, and their members have been specifically trained to collect the desired data.

Collecting environmental data is perhaps the most confusing and difficult area of data recording. The reason for this confusion appears to be that archaeologists have only poorly developed theoretical notions about the relationships between aspects of site location and the environment. The prevailing tactic seems to be to

use as many environmental variables as possible, in the hope that something useful will fall out. The result has been a mushrooming of the quantity of environmental data being recorded. Less than a decade ago, one-or two-page survey forms were the norm. Today many forms are standardized at the state or institutional level, and often they are 10–15 pages long and accompanied by a 20- to 30-page instruction manual!

It is not at all clear that the recent trend toward standardization has either improved the accuracy of data recording or provided the desired data. This is especially true of features of the environment that are difficult to distinguish without quantification, such as plant communities, or those that are known to change through time (e.g., vegetative zones). Instead of having archaeologists, who may be poorly trained, make many of these observations, a more efficient approach would be to determine for a specific project the environmental factors associated with site location (either through background research or in a pilot study) and then train crews to make only these critical observations. This approach encourages flexibility rather than standardization in recording procedures. It may be true that at some later date an archaeologist may find that data pertinent to a specific problem were not collected. But the same thing can happen even under the alternative approach of trying to record everything at once, and worse, the "record it all" approach probably increases the chance that whatever data are recorded are recorded inaccurately.

Once we have decided what data to record, we then need to decide how to record them. Some types of data can only be observed and recorded in the field (e.g., site location, artifact assemblage, etc.), but others, such as vegetation and slope, might be recorded equally well either in the field or in the laboratory. There is no question that data collected in the laboratory are less expensive to acquire and easier for others to replicate than field-recorded data. Before the decision is made to collect data in the laboratory, however, the researcher must determine that the resulting information will be sufficiently accurate and precise. Specifically, if information is going to be taken from 7.5-minute USGS quadrangles, the adequacy of these maps for providing the data at the required scale must be tested, rather than assumed. Verification of test information taken from maps should be carried out before the research design is finalized, and the test data should be selected from a variety of environmental settings.

The decision concerning whether to collect certain data in rhe field or in the laboratory will also be affected by several project-specific considerations. If field crew members do not have the training to recognize vegetation patterns or to distinguish different artifact types, it may be unrealistic to expect them to record such information in the field. On the other hand, there may be instances where variables exhibit interaction effects, making it necessary to record the data for these variables in the field—information that would otherwise be collected in the laboratory. As a case in point, if site size falls below a particular threshold, it might be desirable to record some aspects of microtopography in the immediate vicinity of the site during the field visit. If laboratory determinations of slope rely on calcula-

tions using relatively small scale topographic maps, the resultant data may reflect only a general average in the neighborhood of the site.

The third class of information needed for predictive modeling involves recording of cultural phenomena. In general terms we want to know as much as possible about the activities that took place at a locale and about the timing of those activities. Data pertinent to these objectives describe the nature of artifacts and features present and their spatial distribution. Many of the issues that were discussed in the context of areal survey have correlates at the level of site recording. For example, just as the spacing between surveyors is the single most important factor in determining the number and type of sites found in a survey, the spacing between surface collectors is the primary determinant of the number of types of artifacts collected (or observed) at a site. Questions of sample size and fraction, unit size and shape, and sample design must also be resolved at the site level.

There are, however, fundamental differences between regional survey and site collection strategies. At the regional level we begin with a clearly defined sample universe. At the site level, the first issue to be decided is the boundary of the unit. In areas of high surface visibility, determining the areal extent of a site may not be problematic, in which case defining an appropriate collection or observation strategy is relatively straightforward. At sites with minimal or no surface expression, much of the time spent recording the site will be devoted to defining the boundary, with little or no attempt being made to obtain a representative sample of the cultural assemblage.

A second difference is that at the site level we are sometimes in the position of being able to define the entire population of surface artifacts, or at least a very high proportion thereof. This is especially true of low-density artifact scatters. Often it is less time consuming to flag and map each artifact in the entire site area, and collect them if possible, than it is to grid the site and sample it. Further, because one of the major problems in predictive modeling is site-class definition, and especially functional definition of undiagnostic artifact scatters, complete distributional assemblage analysis is often a requirement rather than a luxury.

In contrast, sites with high artifact densities will probably have to be sampled. These sites are not likely to present major definitional problems, however, since they will usually yield diagnostic temporal and/or functional data. Any type of probabilistic sampling design that ensures that all areas of the site are inspected is likely to yield the data necessary for site-class definition.

Another type of site, the large, low-density artifact scatter, is much more troublesome. In many cases the designation of such a phenomenon as a "site" is a misnomer, if *site* is taken to mean anything other than a defined area of cultural materials. These sites are usually interpreted as resulting from multiple occupations at which similar (or dissimilar) activities may have been conducted. If we are to have any hope of disentangling these multiple occupations, precise distributional information from large block units must be collected. Thus, the grain size of a grid placed over such a site must be at least as large as one cluster of artifacts and features.

Decisions about appropriate survey unit size and shape should be based on a preliminary reconnaissance of the site. Once a grid has been established over the site, an appropriate number of survey units can be selected for sampling, with artifacts and features in each selected unit mapped and collected or observed and recorded.

In some areas of widespread, low-density artifact scatters it is impossible even to distinguish where one site ends and another starts. Van Tries Button (personal communication, 1986), faced with such a situation in the San Luis Valley of Colorado, developed a survey procedure, termed *transect recording*, in which the location, length, and orientation of sets of 2 m wide transects were specified. Transects were spaced every 100 ft and provenienced to a 0.10 mi^2 unit. Counts on all artifacts and on a specified list of environmental attributes found in each transect were made and computer coded. In this way an entire 20,000-acre parcel was surveyed. This approach was highly successful in this case because the entire area could be considered one large, low-density scatter. By not forcing the results into an inappropriate concept (i.e., sites), the researchers were able to make useful statements about the quantity and nature of cultural resources in a reliable and replicable manner.

DATA PROCESSING

The collection and processing of new data for predictive modeling, whether in the field or in the laboratory, has traditionally been a labor-intensive and largely inefficient process. The advent of computers held out the promise that the process of getting information from the field into a form that could be analyzed could be greatly speeded up and streamlined. During the 1960s and most of the 1970s many projects utilized large mainframe computers for this purpose, with varying degrees of success. Yet it was not until the rise of relatively inexpensive microcomputers and associated hardware and software that the potential of automated data processing came within the reach of the vast majority of archaeologists.

It is not our purpose here to review this rapidly changing field. Instead, we will discuss some of the factors that should be considered by those who wish to automate data collection and processing.

Preliminary Considerations

The process of collecting and recording data for predictive modeling should be carefully planned from the beginning of the project. As Sarasan (1981:48) has pointed out, once the research design has been selected and most of the data for a project have been collected, restructuring of the data system may be extremely time consuming, costly, or both, and in certain situations it may indeed be impossible.

In-Field Data Recording Options

At the present time many different system options designed to convert raw data into machine-readable form exist. Not all of those available are suitable for use in field-recording situations, however, and some that are adaptable to field use are more practical in certain settings than others. Factors other than intended location of use also influence the choice of an optimal data-recording system. One of the most important considerations is to minimize the number of steps between data observation and machine-readable record, since this reduces the opportunities for transcription errors (Gaines and Gaines 1980; Nagle and Wilcox 1982). Considerations affecting decisions about data recording will differ between field and laboratory settings of a single project, and data recording will probably be subject to different constraints during each new investigation.

The most commonly used recording format is the familiar, handwritten data code sheet, which is used in various permutations for coding site survey or artifact data. Data code sheets have been in use for a long time and are not likely to be easily supplanted as the primary archaeological tool for field data entry. Handwritten forms are highly portable, survive all but the most adverse field conditions, and provide a readily accessible hard copy of the information they contain. "When all else fails, one can always go back to the field notes" is perhaps the most commonly held (if not always the most accurate) archaeological perception of data recording. On the other hand, most code sheets filled in by hand are not machine-readable and must go through a secondary transcription to attain this state, a step that has the potential for introducing errors into the data.

Nevertheless, variations of the handwritten data code sheet will continue to be used in gathering data, as they should for small to moderately sized projects. Because site survey and/or artifact forms have to be transcribed, they should be designed to follow as closely as possible the intended flow of later machine entry. Chenhall (1975) lists many "do's" and "don't's" for those who anticipate developing and using hand-completed forms as the first stage in data entry.

Another well-known paper format, the optical mark, OMR, or OPSCAN form, possesses many of the advantages of the handwritten data code sheet but is directly machine readable as well (Nagle and Wilcox 1982). Customized forms have been employed to create artifact records in the field (Nagle and Wilcox 1982), to code faunal data (Bonnichsen and Sanger 1977), and to capture site data on several archaeological surveys (Klinger 1977, cited in Schiffer et al. 1978:14; Scholtz and Million 1981:18). If creatively designed, customized OPSCAN forms represent a viable alternative to the use of handwritten code sheets for field data entry since they are well suited to handling interval-scale data as well as other numerically codable, ordinal- and nominal-scale variables in common use in predictive modeling. OPSCAN forms might also be chosen as a means of data entry when poor field environmental conditions eliminate or restrict the use of other automated possibilities.

One of the most promising areas of automation in field data collection is the continuing development of portable data collectors. These machines, often no larger than a standard calculator, record and store data in a machine-readable format that can subsequently be transferred to a more powerful and less portable machine. Portable data collectors, or PDCs, have been used since the early 1970s in the fields of forestry and mining (Cooney 1985). Many of the early PDCs had dedicated functions, such as determining tree height or board feet, which restricted their use to one discipline.

In the late 1970s a number of archaeologists began experimenting with the use of PDCs in field situations (Altschul and Sanders 1984; Stephen and Craig 1984). While the technique was promising, these researchers ran into a number of common obstacles: most notably, excessive power demands, programming difficulties (many of the early machines, such as the Hewlett-Packard 41 series, could only be programmed in a language specific to that machine), storage limitations, communication problems, and the inability to produce paper copy in the field. With the advent of lap or notebook computers, virtually all of these problems have been solved. Computers are now readily available that can easily be carried into the field, are battery powered, have built-in communication capabilities, and can utilize one or more high-level programming languages. Further, the development of battery-powered peripherals, such as microcassette drives and printers, provide the necessary storage requirements demanded by archaeological field situations as well as the capabilities to produce on-site hard copy of field forms and bit-mapped drawings.

Laboratory Data Recording

Although much has been made of the potential use of microcomputers on-site, this is rarely feasible. By nature, surveys are mobile, and microcomputers (even the so-called portables) are ill-designed for this purpose. Microcomputers may be more useful on-site during excavations, but by and large the primary purpose of having a machine in the field is to record and store data, and in this role microcomputers (even with all their power and capabilities) are simply no match for the lighter, less expensive, and more maneuverable PDCs.

Where microcomputers *can* be used effectively is in the laboratory. Here the computer can provide data entry, data storage and management, text editing, and statistical manipulations and can serve as a mechanism to communicate with and transport data to and from other micro- and mainframe computers. For survey projects of even moderate size, data management with a data base management system is probably a cost-effective strategy. As most archaeologists are familiar with these computer capabilities, they will not be discussed further.

We would like to note, however, that all commercial statistical software packages (whether for a micro- or mainframe computer) with which we are familiar require input data to be in the form of a sequential (and generally ASCII format) file. Records in such a file (so-called flat files) usually correspond to a survey unit, a site,

or an artifact, although these record types would not be interspersed on a single file. Since this will probably be the file format in which the vast majority of predictive modeling analyses are conducted, anyone contemplating using a generalized data base management program to store and manipulate his or her data should be cognizant of the fact that it will be necessary to convert the records, or a subset of the records, to a flat file format prior to conducting statistical analyses. Fortunately, most software packages incorporate utility programs to accomplish this step easily, but the capability of creating a sequential output file structure should still be ascertained in advance of selecting any particular software for data base management.

CONCLUSIONS

This chapter has presented an outline of the data needed to create a predictive model, some of the factors that should guide the development of a survey strategy to obtain those data, and the constraints of data collection in a cultural resource management context. By virtue of the fact that different types of data are needed at different times to build a predictive model, the process lends itself very well to multistep survey designs. While each situation will call for a distinct strategy, some general guidelines can be suggested.

The first step of fieldwork should concentrate on three topics: *(a)* magnet sites, *(b)* depositional and postdepositional processes, and *(c)* estimates of site density and of the range of site types. Some sort of informed probing of specific locations (i.e., using information from local informants or regional knowledge) combined with extensive areal coverage (either through imagery or actual flyover) should detect a large proportion of the magnet sites. A detailed geoarchaeological analysis should provide the necessary information on paleo land surfaces as well as indicate past trends in environmental conditions. Finally, some type of small-scale probability sample survey can be used to calculate working density estimates and to obtain some notion of the range in variability in site types. Sample universes should conform to natural units, and the area to be surveyed should be stratified if previous information can lead to the definition of justifiable strata; otherwise, a simple random sampling approach is advisable. The level of survey intensity for this first stage should probably be high.

The second step of fieldwork should be devoted to obtaining the specific information needed to develop the predictive model. Data must be gathered on the relationship between site locations and environmental features and between sites and other sites. Based on the preliminary density estimates and the location of magnet sites, the sample universe(s) should be stratified if at all possible. For example, catchment zones can be defined around each intrinsically important site and treated as separate strata, as can environmental zones that show wide ranges in site density. Optimal allocation formulas can be used to maximize survey resources. It may be necessary to increase the grain size of the grid during this stage of the

survey. This will especially be true if interest focuses on intersite relationships so that a high proportion of survey units must contain two or more sites.

A third step of fieldwork is necessary to test the model. At this stage some form of purposive selection may be used to designate for survey areas predicted to contain sites and areas predicted not to contain sites. Alternatively, the region could be stratified on the basis of high, medium, and low probability of site location, and each resulting stratum could be sampled according to some probabilistic design. Also at this time the geomorphic map of the survey area should be tested. One approach would be to place subsurface tests, such as deep cores, test pits, or shovel probes, according to some multistep sampling design. A second possibility would be to use some type of subsurface test, such as backhoe trenches at specific locales along an alluvial terrace.

Discussions of multistep survey designs are not new in archaeology (e.g., Binford 1964; Judge et al. 1975; S. Plog et al. 1978; Schiffer and Wells 1982; Schiffer et al. 1978). Implementation of such designs, however, is less common, and multistep surveys are almost nonexistent within cultural resource management contexts. By its very nature predictive modeling is a multifaceted process; it is important, therefore, that surveys designed to collect data for predictive modeling projects be multistep as well.

REFERENCES CITED

Altschul, Jeffrey H.

1984 *Bug Hill: Excavation of a Multicomponent Midden Mound in the Jackfork Valley, Southeast Oklahoma.* Report of Investigation No. 81-1. New World Research, Pollock, Louisiana.

Altschul, Jeffrey H., and John C. Sanders

1984 An Automated Approach to Archaeological Site Mapping and Recording. Ms. on file, Statistical Research, Tucson.

Bell, Colin, and Howard Newby

1974 *Community Studies: An Introduction to the Sociology of the Local Community.* Praeger, New York.

Binford, Lewis R.

1964 A Consideration of Archaeological Research Design. *American Antiquity* 29:425–441.

1968 Review of *Rethinking Archaeology*, by K.C. Chang. *Ethnohistory* 15:422–426.

Bliss, C. I.

1953 Fitting the Negative Binomial Distribution to Biological Data. *Biometrics* 9:176–200.

Bonnichsen, Robson, and David Sanger

1977 Integrating Faunal Analysis. *Canadian Journal of Archaeology* 1:109-133.

Butzer, Karl W.

1971 *Environment and Archaeology: An Introduction to Pleistocene Geology.* Aldine, Chicago.

1982 *Archaeology as Human Ecology: Method and Theory for a Contextual Approach.* Cambridge University Press, Cambridge, England.

Chang, K. C.
1967 *Rethinking Archaeology*. Random House, New York.

1968 *Settlement Archaeology*. National Press, Palo Alto.

Chenhall, Robert G.
1975 *Museum Cataloging in the Computer Age*. American Association for State and Local History, Nashville.

Clarke, David L.
1968 *Analytical Archaeology*. Methuen, London.

Clarke, David L., editor
1977 *Spatial Archaeology*. Academic Press, New York.

Cliff, A. D., and J. K. Ord
1973 *Spatial Autocorrelation*. London.

Cochran, William G.
1977 *Sampling Techniques*, 3rd ed. Wiley and Sons, New York.

Cooney, Timothy M.
1985 Portable Data Collectors, and How They're Becoming Useful. *Journal of Forestry* 83:18–23.

Cowgill, George L.
1975 A Selection of Samplers: Comments on Archaeo-Statistics. In *Sampling in Archaeology*, edited by James W. Mueller, pp. 258–274. University of Arizona Press, Tucson.

Dacey, Michael F.
1964 Modified Poisson Probability Law for Point Patterns More Regular than Random. *Annals of the Association of American Geographers* 54:559–565.

Davidson D. A., and M. L. Shackley, editors
1976 *Geo-Archaeology: Earth Science Past and Present*. Duckworth, London.

Dixon, C., and B. Leach
1978 *Sampling Methods for Geographical Research*. Concepts and Techniques in Modern Geography No. 17. Geo Abstracts, Norwich, England.

Doelle, William H.
1976 *Desert Resources and Hohokam Subsistence: The Conoco Florence Project*. Archaeological Series No. 103. Arizona State Museum, Tucson.

1977 A Multiple Survey Strategy for Cultural Resource Management Studies. In *Conservation Archaeology: A Guide for Cultural Resource Management Studies*, edited by Michael B. Schiffer and George J. Gumerman, pp. 201–209. Academic Press, New York.

Dunnell, Robert C., and William S. Dancey
1983 The Siteless Survey: A Regional Scale Data Collection Strategy. In *Advances in Archaeological Method and Theory*, vol. 6, edited by Michael B. Schiffer, pp. 267–287. Academic Press, New York.

Durand, Stephan R., and Jonathan O. Davis
1985 Archaeological Computer Application with an IBM-PC. *Advances in Computer Archaeology* 2:21–38.

Ebert, James I., LuAnn Wandsnider, and Signa Larralde
1984 Theoretical, Methodological and Economic Aspects of Nonsite Surface Survey, Nonsite Sampling and Predictive Modeling. Paper presented at the 49th Annual Meeting of the Society for American Archaeology, Portland.

Flannery, Kent V., editor
1976 *The Early Mesoamerican Village*. Academic Press, New York.

Gaines, Sylvia W., and Warren M. Gaines
1980 Future Trends in Computer Applications. *American Antiquity* 45:462–471.

Galeski, Boguslaw
1972 *Basic Concepts of Rural Sociology*, translated by H. C. Stevens. Manchester University Press, Manchester, England.

Goodenough, Ward H.
1966 *Property, Kin, and Community of Truk.* Archon Books, Hamden, Connecticut.

Goodyear, Albert C.
1975 *Hecla II and III: An Interpretive Study of Archaeological Remains from the Lakeshore Project, Papago Reservation, South Central Arizona.* Anthropological Research Papers No. 9. Arizona State University, Tempe.

Harvey, David
1967 Models in the Evolution of Spatial Patterns in Human Geography. In *Models in Geography*, edited by Richard J. Chorley and Peter Haggett, pp. 549–608. Methuen, London.

Hassan, F. A.
1979 Geoarchaeology: The Geologist and Archaeology. *American Antiquity* 44:267–270.

Hauck, F. R.
1979a *Cultural Resource Evaluation in Central Utah, 1977.* Cultural Resource Series No. 3. Bureau of Land Management, Salt Lake City.

1979b *Cultural Resource Evaluation in South Central Utah, 1977–1978.* Cultural Resource Series No. 4. Bureau of Land Management, Salt Lake City.

Hayes, William L., and Robert L. Winkler
1971 *Statistics: Probability, Inference, and Decision.* Holt, Rinehart and Winston, New York.

Haynes, C. Vance, Jr.
1968 Geochronology of Late-Quaternary Alluvium. In *Means of Correlation of Quaternary Successions*, edited by R. B. Morrison and H. E. Wright, Jr., pp. 581–631. Proceedings of the VII INQUA Congress, vol. 8. University of Utah Press, Salt Lake City.

Haynes, C. Vance, Jr., and G. A. Agogino
1966 Prehistoric Springs and Geochronology of the Clovis Site, New Mexico. *American Antiquity* 31:812–821.

Hodder, Ian
1977 Some New Directions in the Spatial Analysis of Archaeological Data at the Regional Scale. In *Spatial Archaeology*, edited by David L. Clarke, pp. 223–352. Academic Press, New York.

Hodder, Ian, and Clive Orton
1976 *Spatial Analysis in Archaeology.* Cambridge University Press, London.

Hudson, John C.
1969 A Location Theory for Rural Settlement. *Annals of the Association of American Geographers* 59:365–381.

Jacobsen, T., and R. M. Adams
1958 Salt and Silt in Ancient Mesopotamia Agriculture. *Science* 128:1251–1258.

Judge, W. James
1981 Transect Sampling in Chaco Canyon—Evaluation of a Survey Technique. In *Archaeological Surveys of Chaco Canyon, New Mexico*, by Alden C. Hayes, David M. Brugge, and W. James Judge, pp. 107–137. Publications in Archaeology 18A. National Park Service, Washington, D.C.

Judge, W. James, James I. Ebert, and Robert K. Hitchcock
1975 Sampling in Regional Archaeological Survey. In *Sampling in Archaeology*, edited by James W. Mueller, pp. 82–123. University of Arizona Press, Tucson.

King, Leslie J.
1969 *Statistical Analysis in Geography.* Prentice-Hall, Englewood Cliffs.

Kintigh, Keith W., and Albert J. Ammerman
1982 Heuristic Approaches to Spatial Analysis in Archaeology. *American Antiquity* 47:31–63.

Klinger, Timothy C., assembler
1977 *New Hope: An Archaeological Assessment of a Proposed Strip Mine Tract in the Gulf Coastal Plain of Southwest Arkansas.* Arkansas Archaeological Survey, Fayetteville.

Krakker, J., M. Shott, and P. Welch
1983 Design and Evaluation of Shovel-Test Sampling in Regional Archaeological Survey. *Journal of Field Archaeology* 10:469–480.

Kvamme, Kenneth L.
1983 A Manual for Predictive Site Location Models: Examples for the Grand Junction District, Colorado. Draft submitted to the Bureau of Land Management. Grand Junction District, Colorado.

Larralde, Signa, and Susan M. Chandler
1981 *Archaeological Inventory in the Seep Ridge Cultural Study Tract, Uinta County, Utah, with a Regional Predictive Model for Site Location.* Utah Cultural Resource Series No. 5. Bureau of Land Management, Salt Lake City.

Leach, Edmund R.
1961 *Pul Eliya: A Village in Ceylon.* Cambridge University Press, London.

Lightfoot, Kent G.
1986 Regional Surveys in the Eastern United States: The Strengths and Weaknesses of Implementing Subsurface Testing Programs. *American Antiquity* 51:458–504.

Martin, Paul S., and Richard G. Klein, editors
1984 *Quaternary Extinctions: A Prehistoric Revolution.* University of Arizona Press, Tucson.

Matson, Richard G., and William D. Lipe
1975 Regional Sampling: A Case Study of Cedar Mesa. In *Sampling in Archaeology*, edited by James W. Mueller, pp. 124–143. University of Arizona Press, Tucson.

Mayer-Oakes, W. J., and R. J. Nash
1964 Archeological Research Design—A Critique. Paper presented at the 63rd Annual Meeting of the American Anthropological Association.

McManamon, F.
1984 Discovering Sites Unseen. In *Advances in Archaeological Method and Theory*, vol. 7, edited by Michael B. Schiffer, pp. 223–292. Academic Press, New York.

Millon, Rene
1972 *The Teotihuacan Map, Part One: Text.* Urbanization at Teotihuacan, Mexico, vol. 1. University of Texas Press, Austin.

Mueller, James W.
1974 *The Use of Sampling in Archaeological Survey.* Memoirs of the Society for American Archaeology No. 28.

1975 Archaeological Research as Cluster Sampling. In *Sampling in Archaeology*, edited by James W. Mueller, pp. 33–41. University of Arizona Press, Tucson.

Mueller, James W., editor
1975 *Sampling in Archaeology.* University of Arizona Press, Tucson.

Nagle, Christopher L., and U. V. Wilcox
1982 Optical Mark Recognition Forms in Data Entry: Some Applications. *Journal of Field Archaeology* 9:538–547.

Nance, Jack D.
1983 Regional Sampling in Archaeological Survey: The Statistical Perspective. In *Advances in Archaeological Method and Theory*, vol. 6, edited by Michael B. Schiffer, pp. 289–356. Academic Press, New York.

Nance, Jack D., and Bruce F. Ball
1986 No Surprises? The Reliability and Validity of Test Pit Sampling. *American Antiquity* 51:457–483.

Neyman, J.
1934 On the Two Different Aspects of Representative Method: The Method of Stratified Sampling and the Method of Purposive Selection. *Journal of the Royal Statistical Society* 97:558–606.

Orton, Clive
1980 *Mathematics in Archaeology*. Cambridge University Press, London.

Phillips, David A., Jr., Linda L. Swann, and Jeffrey H. Altschul
1984 *Prehistory and History of the Upper Gila River, Arizona and New Mexico: An Archaeological Overview*. Western Division Report of Investigation No. 2. New World Research, Tucson.

Plog, Fred T.
1981 *Managing Archaeology: A Background Document for Cultural Resource Management on the Apache-Sitgreaves National Forests, Arizona*. Report No. 1. Forest Service, Southwestern Region, Albuquerque.

Plog, Stephen
1976 Relative Efficiencies of Sampling Techniques for Archaeological Surveys. In *The Early Mesoamerican Village*, edited by Kent V. Flannery, pp. 136–158. Academic Press, New York.

1978 Sampling in Archaeological Surveys: A Critique. *American Antiquity* 43:280–285.

Plog, Stephen, Fred Plog, and Walter Wait
1978 Decision Making in Modern Surveys. In *Advances in Archaeological Method and Theory*, vol. 1, edited by Michael B. Schiffer, pp. 383–421. Academic Press, New York.

Read, Dwight W.
1975 Regional Sampling. In *Sampling in Archaeology*, edited by James W. Mueller, pp. 45–60. University of Arizona Press, Tucson.

Redman, Charles A.
1974 *Archaeological Sampling Strategies*. Modules in Anthropology No. 55. Addison-Wesley, New York.

Reed, Alan D., and Susan M. Chandler
1984 A Sample-Oriented Cultural Resource Inventory in Carbon, Emery, and Sanpete Counties, Utah (draft). Nickens and Associates. Submitted to Bureau of Land Management, Contract No. YA-553-CT2-1080. Copies available from Bureau of Land Management, Moab District Office, Moab, Utah.

Rogge, A. E., and T. R. Lincoln
1984 Predicting the Distribution of Archaeological Sites: A Case Study from the Central Arizona Project. Paper presented at the 49th Annual Meeting of the Society for American Archaeology, Portland.

Sanders, William T.
1965 *The Cultural Ecology of the Teotihuacan Valley*. Department of Sociology and Anthropology, Pennsylvania State University, University Park.

Sanders, William T., Jeffrey R. Parsons, and Robert S. Santley
1979 *The Basin of Mexico: Ecological Processes in the Evolution of a Civilization*. Academic Press, New York.

Sarasan, Lenore
1981 Why Museum Computer Projects Fail. *Museum News* January/February:40–49.

Saucier, Roger T.
1974 *Quaternary Geology of the Lower Mississippi Valley.* Research Series No. 6. Arkansas Archaeological Survey, Fayetteville.

Schiffer, Michael B., and George J. Gumerman, editors
1977 *Conservation Archaeology: A Guide for Cultural Resource Management Studies.* Academic Press, New York.

Schiffer, Michael B., and S. Wells
1982 Archaeological Surveys: Past and Future. In *Hohokam and Patayan: Prehistory of Southwestern Arizona,* edited by R. H. McGuire and M. B. Schiffer, pp. 345–383. Academic Press, New York.

Schiffer, Michael B., Alan P. Sullivan, and Timothy C. Klinger
1978 The Design of Archaeological Surveys. *World Archaeology* 10(1):1–28.

Scholtz, Sandra C., and Michael G. Million
1981 A Management Information System for Archaeological Resources. In *Data Bank Applications in Archaeology,* edited by Sylvia W. Gaines, pp. 15–26. University of Arizona Press, Tucson.

Stephen, David V. M., and Douglas B. Craig
1984 Recovering Their Past Bit by Bit with Microcomputers. *Archaeology* 37(4):20–26.

Teague, Lynn S., and Patricia L. Crown, editors
1983 *Specialized Activity Sites.* Hohokam Archaeology Along the Salt-Gila Aqueduct, Central Arizona Project, vol. 3. Archaeological Series No. 150. University of Arizona, Tucson.

Thomas, David Hurst
1975 Nonsite Sampling in Archaeology: Up the Creek without a Site? In *Sampling in Archaeology,* edited by James W. Mueller, pp. 61–81. University of Arizona Press, Tucson.

Thomas, Prentice M., Jr., Carol S. Weed, L. Janice Campbell, and Jeffrey H. Altschul
1981 *The Central Coal II Project: A Class II Inventory of Selected Portions of Carbon, Emory, and Sevier Counties, Utah.* Report of Investigation No. 25. New World Research, Pollock, Louisiana.

Tipps, Betsy L.
1984 *The Tar Sands Project: Cultural Resource Inventory and Predictive Modeling in Central and Southern Utah.* P-III Associates, Cultural Resources Report 405-1-8401. Submitted to Bureau of Land Management, Contract No. YA551-CT3-340038. Copies available from Bureau of Land Management, Richfield District Office, Richfield, Utah.

Wobst, H. M.
1983 We Can't See the Forest for the Trees: Sampling and the Shapes of Archaeological Distributions. In *Archaeological Hammers and Theories,* edited by J. A. Moore and A. S. Keene, pp. 37–85. Academic Press, New York.

Wood, John J.
1971 Fitting Discrete Probability Distributions to Prehistoric Settlement Patterns. In *The Distribution of Prehistoric Population Aggregates,* edited by G.J. Gumerman, pp. 63–82. Anthropological Reports No. 7. Prescott College Press, Prescott, Arizona.

Chapter 7

USING EXISTING ARCHAEOLOGICAL SURVEY DATA FOR MODEL BUILDING

Kenneth L. Kvamme

This chapter examines the use of existing archaeological survey data for the development of archaeological locational models. Observe that if an a priori deductive modeling strategy is being pursued, then there is no need for site survey data of any kind for model development (since presumably the "rules" of prehistoric site placement will be derived through theoretical or other means). Hence, this chapter necessarily is oriented toward quantitative model development based on patterns exhibited by empirical data, in this case existing site survey data.

A fundamental assumption made throughout this chapter, unless otherwise stated, is that the *archaeological site* is the basic unit of analysis. For some strategies, a grid cell of small size (e.g., 50 by 50 m) that contains a site or a significant amount of prehistoric cultural evidence is the unit of analysis, but this grid cell type of unit can be assumed to be included in discussions using the site concept.

Our primary concerns when using existing site survey data are with locational and site content information because these two types of data are impossible to obtain without additional survey. We are interested in the *locations* of known sites because most empirical modeling strategies are based on patterns identified in various characteristics of site locations. We are interested in *site content* information for clues that might suggest site function or type, cultural affiliation, or period of occupation. These data are important because we want ideally to develop models for specific types or period groupings of sites. As noted in Chapter 8, however, trustworthy inferences about site function often are difficult to make based on site survey information, and for many sites all that can be said is that a prehistoric site is present at some location.

A third type of information that usually is available in existing site survey reports includes various environmental descriptions pertaining to a site's situation (e.g., vegetation, soils, landform). Although environmental data usually are the very information that is needed for many modeling strategies, the kinds of environmental data commonly included with most site reports often will not coincide with the data requirements of a locational analysis and modeling strategy, and in any

case, the environmental observations usually are inconsistently recorded from site to site. Fortunately, the environmental information reported in existing site survey data is not critical to locational modeling because such data can be observed and measured (and consistently and reliably measured) in virtually any manner on various kinds of maps, aerial photographs, or even through remote-sensing or computer-based geographic information systems techniques (see Chapters 9 and 10).

Collectively, existing site survey data form a large and underutilized body of information that is available in almost any region of study. This body of data represents the cumulative effort of, perhaps, decades of archaeological work performed at considerable cost. Although archaeologists might argue that random samples of site survey data (collected on the basis of regional probabilistic sampling designs) are necessary to make valid regionwide generalizations, new surveys are expensive. Moreover, such an argument neglects an important source of potentially abundant and useful information in the form of existing site survey data. It could be that existing data are well distributed throughout a region of study and are "approximately representative" of a region's archaeology. Alternatively, using procedures discussed in this chapter, it might be possible to make existing data better represent the archaeology of a region through removal or reduction of apparent biases. If existing site survey data could be used in locational studies in place of new survey data, considerable savings in time and cost could be realized.

Of course, the quality of existing site survey data might be questionable and biases might exist in those data. A major focus of this chapter is on ways of removing, or at least reducing, apparent biases from existing data bases in order to obtain better-quality analysis data sets for use in model development or testing. There is no procedure that can correct all biases, of course, and it certainly is not possible to make good data out of bad, but a number of procedures are available that can be used in an effort to reduce certain biases. In most cases, existing archaeological data bases do not constitute a representative sample of the archaeological remains in a region of interest; even in cases where some type of random sample survey results are available, the procedures discussed in this chapter will be useful for preparing other available data for use as, among other things, a test sample with which to assess the performance of site-location models independently (see "Assessing Model Performance" and "Independent Tests," Chapter 8). Problems in the use of existing data are myriad, and only a few can be discussed in detail here. The following pages consider the implications of these problems for model building with existing data. (The statistical and mathematical details of model development are discussed in Chapters 5 and 6; the application of these methods in model development and testing is illustrated in Chapter 8.)

USE OF EXISTING DATA FOR SITE-LOCATION MODELS

A few years ago I conducted a large survey designed to yield a random sample of prehistoric sites that was to be used for developing archaeological models of site

location for the region studied. After the survey was completed I had the opportunity to meet with a statistical consulting group in a university mathematics department. I presented maps illustrating our random sampling design and the locations of the sites that we had discovered. The same maps also happened to show the locations of a few hundred sites known to exist prior to the survey. Although we discussed several interesting topics, the one that struck me most forcibly was that the statisticians were amazed that I had conducted such a large and expensive survey when several hundred site locations were already known for that region.

This reasoning went against all my archaeological training and against what I perceived as an accepted notion in settlement archaeology: that in order to make valid regional inferences about archaeological site location (or any other) patterning one needed representative samples chosen on the basis of probabilistic sampling theory. This position has been stated by Binford:

> Probability sampling is . . . a major methodological improvement which, if executed on all levels of data collection in full recognition of the inherent differences in the nature of observational populations which archaeologists investigate, can result in the production of adequate and representative data useful in the study of cultural process [1964:439].

The statisticians did admit that my sample seemed very nice, but they pointed out that sampling is a pragmatic effort conducted for the purpose of reducing costs. That I had sampled in the first place indicated a concern for cost, yet I conducted an expensive survey even though the previously known sites existed in large numbers and appeared to be well distributed in the region. When I asked them about the role of statistical theory in model development, they suggested that I worry less about theory and more about how well the model works in practice.

In Chapter 8 it is emphasized that from a statistical standpoint *any* procedure—ranging from statistical techniques to simple mathematical rules or even armchair theory—might appropriately be used as a basis for site-location model development. What matters is how well a model works in application, how accurately it performs on future cases. Given this perspective, it is appropriate to use any type of procedure as well as any source of data (such as existing site-file information) in model development. In order to determine how well a model will perform in practice (and here I refer to *any* type of model, including those formulated deductively), independent testing procedures are required, and in this case methods of statistical inference *must* be applied. *Independent testing* means that a model is applied to data independent of the data set used to build the model (note that deductively derived models are not built with data, and therefore any data set is independent of these models), which provides a test of model performance. Statistical theory can then be applied to the test results (if the test data constitute a representative sample) in order to assess the significance of the resulting model performance and construct confidence limits around model accuracy rates. (The reader is referred to the section on assessing model performance in Chapter 8 for a more detailed discussion of these issues and procedures.) The purpose of the current chapter is to

examine various problems that can be encountered in existing archaeological site survey data bases and to recommend ways of correcting some of the more apparent problems in order to obtain data sets better suited for regional analysis and modeling purposes. In other words, this chapter examines methods for reducing obvious biases so that the locational patterns apparent in a final analysis sample of existing archaeological data are more likely to be representative of overall locational patterns within the region of interest.

PROBLEMS AND BIASES IN EXISTING SITE SURVEY DATA

When examining existing archaeological site survey data bases from a region one is often struck by the great variation apparent in the quality of the data. It has been appropriately noted that the greatest source of variability in the archaeological record may be due to the behavior of the archaeologist. This variation stems from a number of factors, ranging from differing standards of quality or practice between different archaeologists to changes through time in accepted field practices to variability in the goals and research plans of individual survey projects.

The ways in which different field projects, archaeologists, and field crews perform fieldwork and define, identify, and record archaeological sites introduce the major sources of variation, bias, and inconsistencies in existing site data bases. Chapter 4 describes some of the operational problems in defining sites on the basis of diffuse scatters of artifacts. Sites defined by one project may not constitute sites by another project's definition. Not only does the lack of standard archaeological procedures, such as field methods and operational definitions of sites, create inconsistencies in the data base, but differences in research goals from project to project, even within a single region, create major inconsistencies in regional data bases.

The problem goes deeper than this, however. Even within a single project, sites mighr, in practice, be defined differently owing to differences in the quality of individual field personnel and crews or because of other factors, such as insect density, adverse weather, terrain roughness, crew tiredness, or the arrival of a Friday afternoon. Budgetary constraints can also influence the quality of data collected when, for example, a contractor has a fixed price contract but site densities are greater than expected; this can lead to "hurrying" the survey. Schiffer and Wells (1982:346) note that "this is accomplished by increasing crew spacing or reducing the recording time. We suspect that such modifications in technique are rather common, if seldom admitted in the final report." These practices, of course, can lower the quality of the resulting data.

Several factors influencing archaeological survey results and the quality of the retrieved data are summarized by Schiffer and Wells (1982). A principal factor is survey intensity or crew spacing. Crew spacing not only affects site discovery rates but also the sizes of discovered sites (Plog et al. 1978). Small sites and cultural features tend to be missed when crew spacing is large (Wandsnider and Ebert 1984).

Narrow spacing, however, dramatically increases survey time and effort and therefore costs (Figure 7.1).

The nature or *obtrusiveness* of the archaeological evidence determines the likelihood that a particular archaeological feature, such as a site or an artifact, will be discovered given a specified level of survey intensity (Schiffer et al. 1978). A mound or architectural feature, for example, has a higher chance of discovery than a single, isolated flake. Low-intensity surveys (those with wide spacing) tend to bias resulting archaeological samples in favor of more obtrusive remains (Schiffer and Wells 1982).

Difficulty of access, a common problem in many regions of the western United States, might mean that samples are biased against difficult-to-reach regions. In regions with relatively few access roads, for example, sampling units might be placed with the restriction that units lie within some maximum distance of an existing road. Even when it is possible to arrive at hard-to-reach places, the limited amount of time left in the day after travel might lower the quality of resulting survey in those regions. Private land ownership presents similar difficulties when landowners refuse access (Schiffer and Gumerman 1977:187). Indeed, in western regions, where most archaeological survey work tends to be conducted on federal or state lands, the lack of comparable site data from private properties presents a

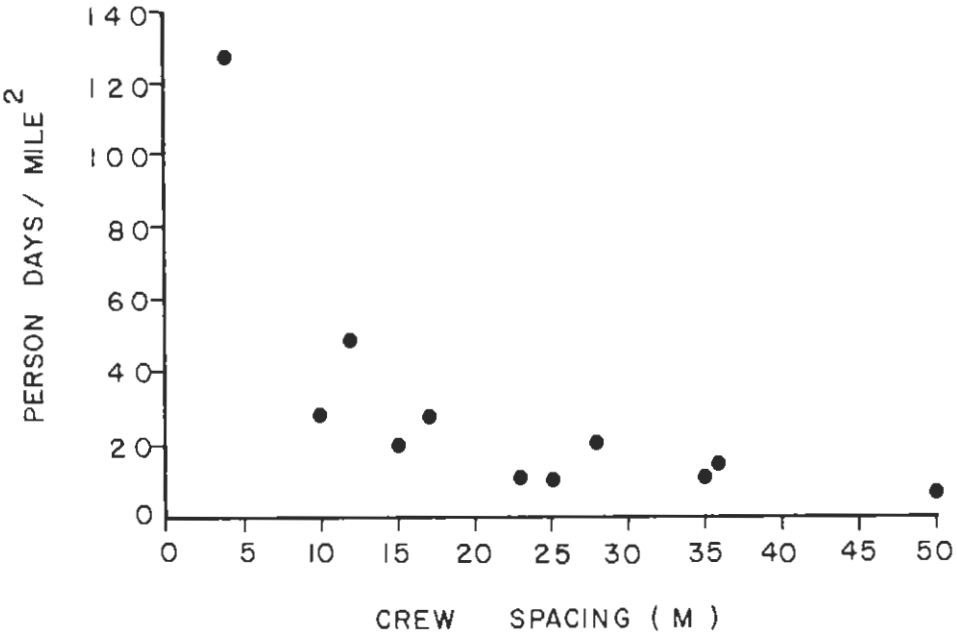


Figure 7.1. Relationship between crew spacing and survey rate (after Schiffer and Wells 1982:354).

severe source of bias to regional archaeological data bases, because private property often includes some of the best agricultural lands as well as the best areas for hunting and plant collecting, and prehistorically, it is these very places that often were the most critical to site placement.

Variable archaeological visibility, due primarily to vegetation cover, introduces another major source of potential bias. Planted fields, swamps, or forests might offer poor visibility and low archaeological discovery rates, while desert regions or sagebrush-grassland settings usually offer high visibility and excellent site discovery rates (Schiffer and Gumerman 1977:187). Study regions containing zones with markedly different levels of visibility are likely to have existing site data bases biased toward the more visible zones.

Perhaps one of the principal weaknesses of existing data bases is that the sum total of previous work in a given region constitutes an unplanned effort. In other words, strong locational biases typically exist in the areas that have been field inspected within a region. For example, early work often was conducted only at the most accessible and visible sites, while much contemporary survey is conducted primarily in areas of planned development. Thus, existing site data may be strongly biased toward certain types of settings and may not constitute a representative sample of sites within a region.

An additional problem is that sites might not be accurately located on maps. For modeling approaches that focus on the specific locations of sites, accurate placement of sites on maps is of critical importance since characteristics of the actual locations, such as environmental properties, are often used as a basis for modeling. In actual field practice it is often difficult to locate oneself precisely, particularly in forested areas with few nearby landmarks. Field crews often get lost or misread maps. Moreover, early archaeological surveys often did not have access to good maps and offered only verbal descriptions, directions, and rough locational sketch maps.

This problem is further compounded as site locations are transferred from map to map. In examining existing site files for one Bureau of Land Management (BLM) study, I found that the original site forms were available as well as the district's master management maps. The latter are a set of maps that can be found in any regional BLM office and contain the most up-to-date information on the locations of all known sites and field-inspected regions. In this BLM district, the majority of the sites were extremely small lithic scatters (essentially points on maps). When the site forms, which included copies of original maps, were compared with the BLM's master maps, many sites were found to have been mislocated when they were copied from the original to the master maps (Figure 7.2). In fact, almost 10 percent were mislocated by more than 100 m (one-sixth of an inch on 1:24,000 scale maps), and several were even placed on the wrong drainage!

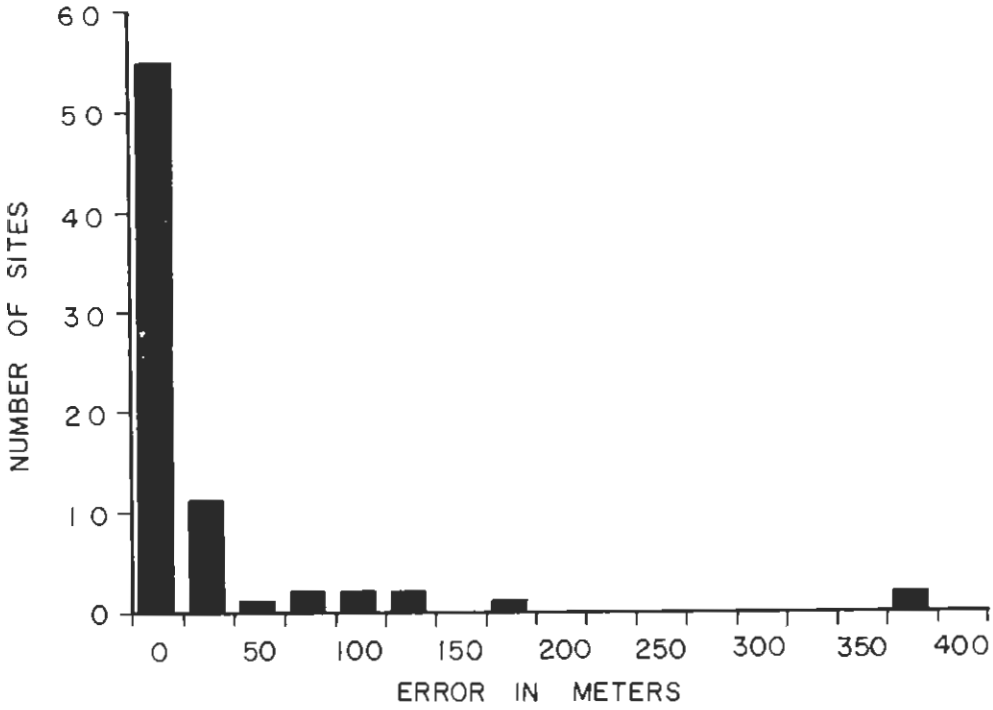


Figure 7.2. Errors in the locations of sites resulting from copying site locations from one map to another.

PROCEDURES FOR REDUCING DEFICIENCIES AND BIASES IN EXISTING DATA

A number of problems with existing data bases were presented in the previous section. In order for researchers to use such data in archaeological model development they need to eliminate data of questionable quality and to reduce the effects of apparent biases.

If possible, the original site forms should be obtained in order to assess the quality of the initial site-recording effort and to eliminate secondary sources of error that might be introduced by later handling of the data by other investigators (as in the example discussed above). Certain minimal standards might be established; precise location of the site on a USGS 7.5-minute map might be required, for example, along with a description of some minimal amount of archaeological evidence. Sites not meeting these standards might be eliminated at this stage.

When a pool of minimum-quality sites has been obtained based on inspection of site forms, it would be prudent, depending on available funds, to examine in the

field a random sample of the sites recorded by each major investigator in the area. This practice would allow verification of locational accuracy on maps as well as assessments of site content and function. It would also be worthwhile to resurvey at high intensity regions that have been field inspected by other researchers in order to obtain data on site discovery rates. These rates might then be used as a means of bias correction through the subsampling or weighted analysis techniques described below.

When use of existing data in site-location model development is considered, bias must be viewed in terms of current modeling goals. For example, a survey conducted for the discovery of only Paleoindian sites is not relevant to a site-location model for Puebloan villages. Similarly, a survey conducted in pine forests does not bear on models for grassland settings.

The nature of bias also must be considered in terms of the type of modeling approach used. Models that examine characteristics observed at the actual locations of sites or models that use a small-size quadrat approach in which characteristics of quadrats with sites are examined are particularly sensitive to the happenstance locational biases of previous surveys. For example, if 60 percent of one part of a study region has been field surveyed but only 20 percent of another part, indiscriminate use of the site data without regard to these survey proportions can bias a resulting model toward characteristics of the more extensively surveyed zones. On the other hand, modeling approaches that partition a region into discrete categories, such as environmental communities, and then project site densities in each community are less sensitive to this factor. In this approach, if one community has been 20 percent surveyed but another 60 percent, so much the better for the latter community, since the resulting estimates of site density would presumably be more reliable because they are based on more information.

Two major approaches might be investigated as a means of reducing the influence of known biases in existing data. *Subsampling* attempts to reduce biases by undersampling areas that have been extensively examined and by oversampling areas that, by comparison, have been little examined. This procedure usually requires that some information be disregarded during the model-building process, but it should be realized that the sites eliminated during this part of the project might be reserved to provide independent tests of locational models at some later point. *Weighted analyses*, on the other hand, permit retention of all information, but the impact of an individual case (e.g., a site) on the analyses can be weighted by, for example, the relative importance of that case relative to other cases (see below).

Subsampling

A common problem in existing site-file data bases is unequal survey coverage in various regions of a study area; these inequalities are a result of the use of nonprobabilistic designs and purposive survey that is commonly required for various forms of cultural resource clearance. Early surveys typically examined only

the most ideal or most easily accessible regions, simply for purposes of site discovery. Unequal survey coverage also occurs among different archaeologists or projects owing to variation in crew spacing, vegetation cover, and other factors. A goal of subsampling is to obtain a subset of the total number of sites available in the entire study area such that many of the regional biasing factors are reduced in the final subset. A number of approaches might be used to accomplish this goal.

One approach that helps to reduce the effects of unequal amounts of survey in different regions of a study area is to divide the area into discrete categories, such as environmental communities, and then to sample each category in a way that will correct for the inequities. A hypothetical study area containing three communities is portrayed in Figure 7.3a. Forty percent of community A has been field inspected, 20 percent of community B, and 60 percent of community C. In developing a model for the entire study area it is important to remove the biasing effects of the more heavily surveyed communities. This might be accomplished by selecting 100 percent of the sites in stratum B for the analysis sample and taking a simple random sample of 50 percent of the sites in stratum A and 33 percent of the sites in stratum C; this would yield a 20 percent overall sample of sites in the study area.

Another subsampling approach attempts to provide an analysis sample with a more uniform distribution of sites from within a study area. It is important to attempt to obtain a regional sample that is well distributed across the area of study in order to ensure that site location variation from throughout the entire region is included in the sample. In this approach a grid may be superimposed over the study area or over each stratum in the study area (Figure 7.3b). Depending on the size and nature of the study area the grid might be as large as a township (6 by 6 mi) or as small as a hectare (100 by 100 m). The analysis sample for the gridded study area or gridded stratum is selected by choosing sites from within each grid unit, which creates a more uniformly distributed sample. For example, let us assume that the gridded region in Figure 7.3b is a portion of environmental community C in Figure 7.3a. A simple random sample of 33 percent of all the sites in Figure 7.3b could, by chance, cause some of the gridded cells that contain sites to contribute no sites to the sample and others to contribute many. If a 33 percent simple random sample of the sites within each grid cell were taken instead, this would help to ensure a better-distributed analysis sample.

A third subsampling approach may be used when large clusters of sites exist in a data base. Clusters of sites can have adverse effects on later analyses because the clustered sites may have highly related characteristics rather than offering new and independent information. A field-inspected region containing a single cluster of many sites along with a number of dispersed sites is portrayed in Figure 7.3c. If a subsample of 20 percent of all sites in the region were randomly selected for an analysis sample, it is likely that all or almost all of the selected sites might be from the single cluster. Yet, multiple sites from the same cluster might yield much redundant locational information, and it might be desirable to incorporate the locational variation of sites *outside* the cluster into the sample when the goal is a regionwide model and most of the region of concern is outside the cluster. This can

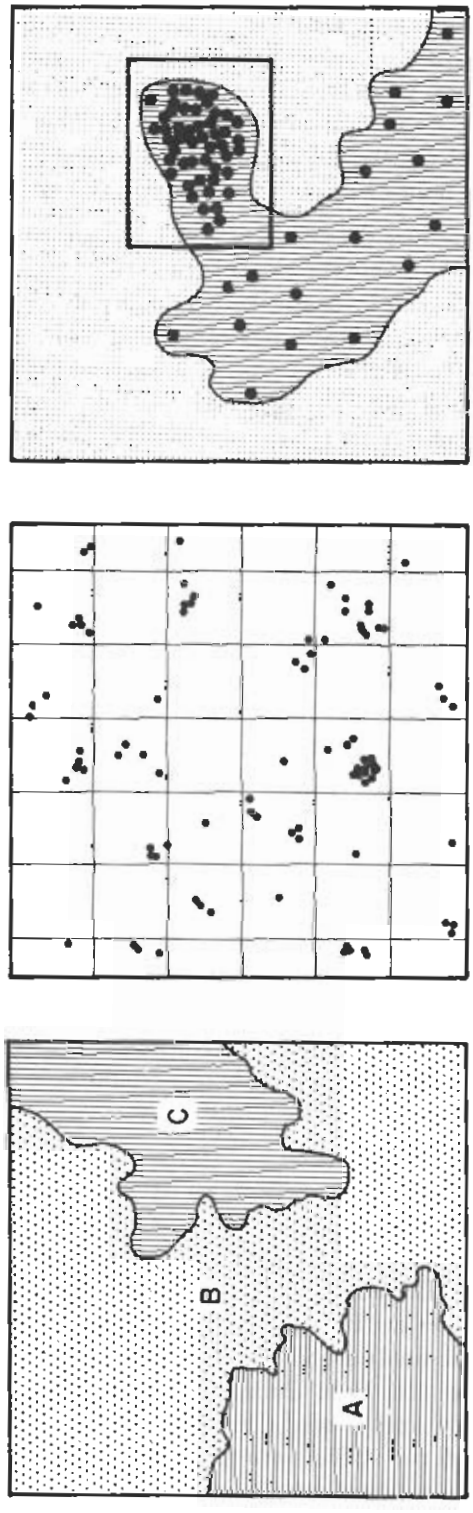


Figure 7.3. Illustrations for bias correction procedures. (A) Three environmental communities with unequal amounts of survey coverage: 40 percent of community A has been surveyed, 20 percent of community B, and 60 percent of community C. (B) A grid superimposed on a region to allow better-distributed samples by selecting sites from each grid cell (dots represent sites). (C) A surveyed region (dark area) containing a cluster of many sites. The small rectangle represents a "cluster" stratum.

be accomplished by stratifying the area into a cluster region and a noncluster region (Figure 7.3c) and taking a simple random sample of 20 percent of the sites in each region.

It might even be desirable, under certain circumstances, to reduce the influence of major clusters still further. This could be accomplished, for example, by taking a larger sample of sites outside denoted clusters (e.g., 30 percent) and a smaller sample of sites within clusters (e.g., 15 percent). The goal might be to develop a model that performs well for the portion of a study region that lies outside clusters. This would be particularly useful where previous investigation has shown that sites from major clusters tend to possess locational properties different from those of sites outside clusters. By taking a smaller sample of sites from clusters, one can reduce the influence of those sites in an analysis. On the other hand, the very presence of clustering can be indicative of desirable locations that need to be included in a sample. Hence, some thought should be given to the goals of the analysis and to the behavioral implications of such patterns when one is using clustered data. The presence of significant clustering can be determined through simple statistical tests described by Clark and Evans (1954), Dacey (1973), and Thomas (1971:41–43).

Weighted Analysis

Weighted analyses can present an alternative to the elimination of data when existing site information is used for model development. Individual cases or sites can be assigned a weight that affects the influence of that site in subsequent analyses. Sites with more "important" location information (e.g., those that lie in undersurveyed regions) can be assigned more weight, and sites with less important location information (e.g., from well-surveyed regions or from major site clusters) can be assigned less weight. In this manner it is possible to utilize information from all or most of the sites, while correcting for certain biases at the same time.

Common statistical analysis computer programs, such as the Statistical Package for the Social Sciences (SPSS 1983), the Statistical Analysis System (SAS Institute 1985), and BMDP Statistical Software (Dixon et al. 1983), allow case weighting as an option for many procedures. A weighted sample mean is given by

$$\bar{x} = \left(\frac{1}{\sum_{i=1}^{n} w_i} \right) \sum_{i=1}^{n} w_i x_i$$

and a weighted variance by

$$s^2 = \left(\frac{1}{\sum_{i=1}^{n} w_i} \right) \sum_{i=1}^{n} w_i (x_i - \bar{x})^2$$

where x_i is the sample value for the ith case (site), and w_i is the weight associated with that case. Note that if $w_i = 1$ for all i cases, these equations reduce to the traditional formulas for mean and variance.

To illustrate how these formulas might be applied, the first problem area of the previous section, environmental communities with disproportionate areas of survey (Figure 7.3a), will be examined. The hypothetical region contains three communities, A, B, and C, of which 40, 20, and 60 percent, respectively, have been surveyed. Suppose, for simplicity, that 4 sites were found in zone A, 3 in zone B, and 6 in zone C, for a total of 13 known sites (Table 7.1). The subsampling approach described in the previous section called for selecting half of the zone A sites and a third of the zone C sites, which would provide an approximated overall 20 percent sample consisting of only seven sites. The weighting approach merely assigns weights to all of the cases such that a site's contribution is inversely proportional to the percentage of area that has been surveyed (Table 7.1). Thus, a site in zone B (of which only 20 percent has been surveyed) carries twice as much weight as a zone A site (of which 40 percent has been surveyed) and three times as much weight as a zone C site (of which 60 percent has been surveyed).

In conducting a site-location analysis encompassing multiple regions, as in Table 7.1 and Figure 7.3a, weighting can permit the archaeologist to emphasize features peculiar to undersurveyed regions. For example, let us say that zone B

TABLE 7.1.

Example of weights applied to data as a means of bias correction

Site	*Slope*	*Distance to Water*	*Stratum*	*Percent Surveyed*	*Weight* (w_i)
1	2	50	A	40	.9286
2	6	70	A	40	.9286
3	3	20	A	40	.9286
4	1	30	A	40	.9286
5	12	180	B	20	1.8571
6	18	350	B	20	1.8571
7	15	290	B	20	1.8571
8	0	60	C	60	.6190
9	2	100	C	60	.6190
10	1	90	C	60	.6190
11	4	50	C	60	.6190
12	3	120	C	60	.6190
13	1	90	C	60	.6190

Normal weighting ($w_i = 1$):

		Slope	Distance to Water
$\bar{x}$	=	5.23	115.38
s^2	=	34.8589	10,110.237

Stratum weighting (using equations given in text):

$\bar{x}$	=	7.81	153.57
s^2	=	25.2627	14,877.342

contains more variable terrain in the form of hills and ridges and fewer sources of water than zones A and C. Hypothetical measurements of slope and distance to nearest water given in Table 7.1 show that, without weighting (i.e., $w_i = 1$), the measurements from the sites in the more heavily surveyed zones A and C dominate, yielding a mean slope of only 5.23 and a mean distance to nearest water of 115.38. When weights giving increased influence to the zone B data are used, however, the weighted mean values exhibit greater slopes (a mean of 7.81) and distances to water (a mean of 153.57), reflecting the greater steepness of hillslopes and the paucity of water in zone B.

The utility of case weighting is not restricted to altering regional survey coverage bias; this procedure can be applied to other sources of bias as well. If reliable estimates can be made of site discovery rates under different types of vegetation cover, the discovered sites in zones offering less visibility might be given greater weight in analysis. A similar approach could potentially be applied to correct for differences in site discovery rates between different archaeologists or projects. The use of weighting to correct any of these forms of bias should be carried out only after thorough consideration of the available evidence, however.

Finally, it is important to note that not only can weighted means and variances be computed, but also covariances, which open the doors to the host of multivariate procedures discussed in Chapters 5 and 8.

EVALUATION OF SITE-LOCATION PATTERNING AND MODEL BUILDING WITH EXISTING DATA

When this stage is reached it must be assumed that the researcher believes he or she has a reasonably good sample of existing sites with which to work. The data might exist in several groups, each corresponding to a different site type. The investigator must decide on the kinds of phenomena that should be investigated for possible relationships with the locations of sites and then devise ways to make these phenomena operational. In other words, the variables that are to be investigated must be defined. An overview of some of the variables commonly used in site-location research and of the ways in which they can be made operational is given in Chapter 8. Once the variables are defined, they must be measured or observed on maps at each of the sample site locations, either by hand (Chapter 8), through remote-sensing techniques (Chapter 9), ot through computer technology using geographic information systems (Chapter 10).

A usual step in the model-building process (e.g., Larralde and Chandler 1981; Thomas and Bettinger 1976) is to examine the data at this point through use of histograms, descriptive statistics, or simple univariate statistical procedutes. In this way it is possible to identify variables that are more likely and less likely to have some bearing on the locations of sites in general or of individual site types.

The empirical data can then be subjected to a variety of modeling approaches ranging from simple mathematical rules to multivariate statistical techniques. A single-class classifier approach (Lin and Minter 1976; Thomas and Bettinger 1976) can be used to model the distribution of individual site classes, or a control-group approach consisting of background environment measurements at locations where sites are *absent* might be used to contrast locations where there are no sites with the locations of known sites using a variety of quantitative classification techniques. A wide range of approaches using a variety of techniques is illustrated in Chapter 8. As noted in that chapter, any form of decision rule may properly be used to develop a modeling procedure for classifying locations—for example, as site-likely, site-type-likely, or site-unlikely locations. Admittedly, some procedures work better than others, and statistical procedures generally work best when the required assumptions are fully met. Once a modeling procedure is developed, however, its performance must be assessed using statistical theory, an independent sample of data (of the kind of site being investigated and from the region being modeled), and a sample that can be argued to be representative of the sites in the region.

Assessing a Model and Determining Additional Data Needs

A fundamental question that must be asked when evaluating a model based on existing data is whether or not the model might be biased. Even if a developed model successfully predicts locational patterns similar to patterns exhibited in the existing site data base, how certain can we be that the existing site data patterns are representative of the locational patterns of as-yet-undiscovered sites in unsurveyed regions? Despite careful data evaluation and crude attempts at bias removal, it is possible that the bulk of existing sites really are not representative of sites in the general study area, and there is no way to determine whether or not this is the case unless some form of data known to be representative of sites in the region at large are obtained with which to test the model.

An initial and simple test of model performance may be obtained simply by applying the model to the same data used to build the model. Although at best this procedure yields an inflated view of the model's true performance, it can provide an immediate indication of model deficiencies. The predictions of site locations made by the model might be categorized along several dimensions to assess performance in a number of areas (Table 7.2; see Chapter 8 for a discussion of the necessity for reviewing model predictions of site-absent locations or nonsites as well). For example, the model might be examined to see how well it predicts various functional or temporal site types or various subtypes of sites (the columns in Table 7.2). Similarly, the performance of the model relative to different environmental settings, such as various plant communities or topographical situations, might be assessed (the rows in Table 7.2). Deficiencies at this stage should be taken seriously; if they exist here they certainly will exist when the model is applied to independent and new samples.

TABLE 7.2.

Assessing model performance along several site-type and environmental categories. In testing a site-location model, the percentage of correct model predictions for each site type are assessed along the columns and the percentage of correct model predictions for each environmental category are assessed along the rows.

	Site Type 1	*Site Type 2*	. . .	*Site Type k*
Environmental Category A				
Environmental Category B				
Environmental Category C				

. . . *Percent Correct*

Site Predictions . . .

Environmental Category *p*				

Model tests that are more independent and can yield a truer picture of actual model performance may also be performed using existing data. One independent test uses sites in the existing data base that were *not* used to construct the model—entries that were eliminated during attempts at bias removal, for example. Such sites represent independent information, and the model can be applied as shown in Table 7.2. A somewhat better approach is called *split sampling* (Mosteller and Tukey 1977:38); with this procedure the analysis data is randomly split into two groups, a model is built with one half, and the remaining data are used to provide an independent test of the model. The jackknife procedure (Mosteller and Tukey 1977:133) presents yet another alternative. In this procedure one case in the analysis data set is temporarily "thrown out," the remaining data are used to build a model, and the single case is used to test the model. This process is repeated using each case in turn to yield an independent assessment of model performance. (Split sampling and jackknifing are discussed in more detail in Chapter 8.)

Since existing data often are highly clustered (Figure 7.3c), the traditional split-sampling and jackknife approaches still might yield an inflated picture of

model performance. A site that is part of a cluster of sites might exhibit characteristics that are highly related to the other sites in the cluster. When that site is used in a split-sample or jackknife procedure, it does not necessarily yield an independent test since its characteristics are related to those of other sites, some of which may have been used to develop the model. An alternative that might offer less inflated results is to superimpose a large grid, like that shown in Figure 7.3b, over the region and to use the grid cells as the basis for the split-sample or jackknife techniques. For split sampling the individual cells are split at random into two groups, and the analysis proceeds with sites in the selected half of the grid cells while sites in the remaining half are reserved for model testing. In the jackknife approach the sites in the kth grid cell are eliminated from the kth model and are then used to test that model independently, with this process repeated for all k cells.

Such testing procedures, however, are only as good as the data to which they are applied, and as mentioned earlier, existing data might inherently be strongly biased. Independent and representative data are therefore needed if we are to assess model performance in a reliable and confident manner.

In many federally administered regions and districts some form of random sampling survey may well have been conducted in the past. These data can be used for model testing if it can be argued that they are representative of sites (or the site type of interest) in the whole region and if the site sample was suitably constructed and sufficiently large. Not only can these data be used to assess accuracy (Table 7.2), but statistical significance can also be determined and confidence limits around the predictions can be calculated. Since the width of a confidence interval is directly a function of sample size, relatively large test samples are desirable. For example, if a model accuracy rate of 80 percent correct is obtained, a sample size of 50 yields a 95 percent confidence interval width of 19.8 percent (±9.9 percent), a sample size of 100 yields an interval width of 15.5 percent (±7.8 percent), and a sample size of 200 yields an interval width of 11.0 percent (±5.5 percent; Hord and Brooner 1976).

Collecting and Integrating New Data in Model Development

During model testing through use of existing data or through use of independent test data it might be discovered that a model underperforms for certain types of sites. Alternatively, a model might perform poorly when applied to certain environmental settings—grassland settings, for example (Table 7.2).

In order to attempt to remedy these failings, the researcher might go back to the existing data base (especially if it contains a number of sites eliminated from the analysis through subsampling or for other reasons) and, using the above examples, attempt to incorporate more grassland sites or more sites of the type being underpredicted. If a weighted analysis approach is being used, the investigator might simply assign more weight to site types or sites in environmental settings that are being poorly modeled. The model-building and model-assessment stages then might be repeated.

Another approach to remedying modeling problems is to develop a specific model for the particular environmental setting or site type that is being incorrectly predicted (Stone 1984). This tactic might be more successful than refinement of the original model, since a site-type or environmentally specific model would only focus on the locational variation exhibited by the particular setting or site type. It should be noted, however, that when analyses become too fine-grained, as when specific site types or environmental communities are investigated, available sample sizes can become prohibitively small.

A last alternative when one is faced with the problems of under- or overprediction by a site-location model is to conduct a new survey designed to obtain more data from deficiently predicted environmental regions or site types. This is a last resort, due to costs, and should be performed only when the researcher is certain that the modeling application warrants collection of new data. It might be that it is not possible to model the locations of sites in a specific environmental community successfully (owing to a low level of patterning with respect to the variables examined, for example) regardless of the amount of data available. The collection of new data in this case would not offer any improvement to the modeling situation. Before initiating a new survey the investigator should consider this possibility by examining the quality and amount of the existing data.

When implementing a survey for the purpose of providing more information about a particular region, such as a specific environmental community, some form of random sampling design should be used. Sites discovered by this survey could then be compared with previously known sites in the same community. This comparison can entail visual inspection of the shapes of histograms of the measured variables, descriptive statistics, indices of difference, and statistical tests for differences, such as the *t*-test. If differences between the samples are found, this would suggest that new and different information might be contained in the new sample. The new data might then be incorporated into the analysis data base or analyzed as a separate data base, and the model-building and testing processes could be reinitiated.

New site data inevitably become available as archaeological work continues in a region. Model updating and testing using these new data can be performed as an ongoing process. The techniques used in evaluating existing data should also be applied to these new data; i.e., the quality of site recording and survey should be investigated, and appropriate bias-removing techniques, such as subsampling to reduce locational survey bias, should be employed.

EXAMPLE ANALYSIS

A settlement pattern study of Mesolithic sites in the Federal Republic of Germany (Kvamme and Jochim 1988; also see Kvamme 1986) will be described here as an example of the use of existing site data as a basis for locational modeling. Although this study does not illustrate many of the bias-reduction techniques discussed above, it does illustrate the locational patterning that can be found, and

the kinds of interpretations of results that can be made, given the biases that might exist in a body of regional archaeological data. This study focused on a region near Stuttgart where there are many recorded Mesolithic sites. The journal *Fundberichte aus Schwaben,* which contains regional archaeological reports of investigations by local amateurs, was used to obtain the locations of 170 known Mesolithic sites in the region. Since the site descriptions were very terse it was not possible to assess quality of reporting, nor was it possible to field check any of the sites. The sites did, however, appear to offer a fairly good spatial distribution that was well spread throughout the 940 km² study area (Figure 7.4a).

Previous research in the Mesolithic of northern Europe had suggested a number of relationships between the physical environment and patterns of settlement. Nine environmental variables were selected for this study (Kvamme and Jochim 1988), largely on the basis of previous work. These variables are elevation, slope, aspect, local relief, a measure of view quality, a measure of shelter potential, horizontal distance to nearest water, vertical distance to water, and horizontal distance to nearest third-order stream (see Chapter 8 for a discussion of how these variables can be defined). Measurements of each variable were made at the locations of the 170 known Mesolithic sites, and the same measurements were made at 100 m intervals across the entire background environment (a total of 84,000 measurements for each variable). The large number of measurements was possible owing to the use of computer-based geographic information system (GIS) techniques (see Chapter 10 for discussion of how the computer approximates measurements on the basis of a regular grid system).

The methodological premise of the study was that, in order to determine significant environmental patterning at site locations, one must contrast empirical data measured at known sites with the same data measured in the background environment. For example, if only the site locations were examined, as is usually the case, the data might indicate a major tendency for south-facing aspects. Such a tendency in the data could reflect a significant pattern, or conversely, the entire study region might generally possess a south-facing orientation, in which case the pattern exhibited by the sites would only be a reflection of the background environment; it is an examination of the background data that allows us to make this assessment. For each variable, the data measured at the 170 sites were contrasted with a representative sample of 3201 measurements taken from the background environment using Student's t-statistics as a rough guide for differences between the two groups. Since the a priori chance of an as-yet-undiscovered Mesolithic site occurring in one of the background samples was assumed to be extremely low, the two classes could be argued to be reasonably distinct, although the representativeness of the Mesolithic sample and the general independence problem of spatial samples forced cautious interpretation of the statistical results.

The analysis results (Table 7.3) indicate a number of strong patterns of contrasts between site locations and the background environment (in the original study, detailed histograms were also examined). The sites show a strong tendency toward level ground slope (Figure 7.4a), for regions of great relief, and for higher

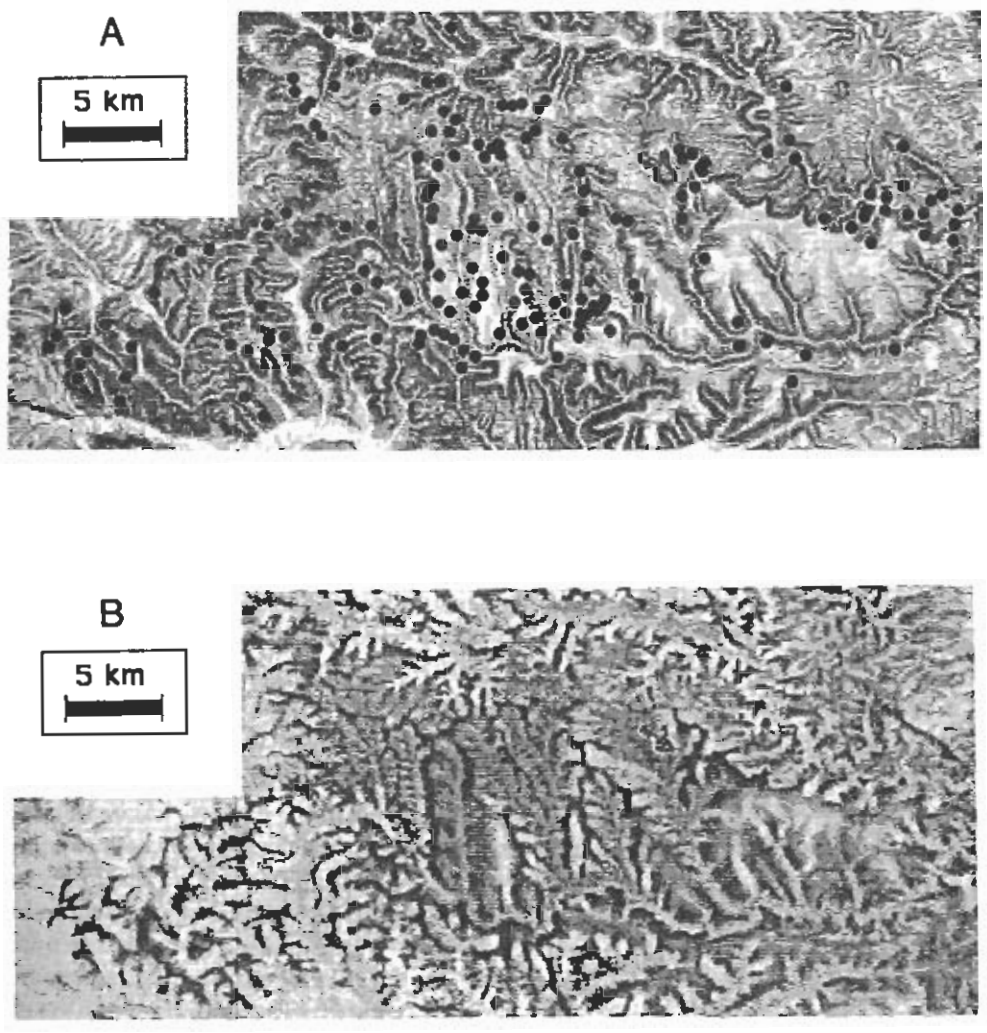


Figure 7.4. GIS-generated images. (A) Locations of recorded Mesolithic sites plotted on a computer-generated image of the study region in southern Germany. The image was obtained by calculating slope every 100 m and shading the image by degree of slope. (B) Image of the multivariate model of Mesolithic site location mapped over the entire study region.

TABLE 7.3.

Descriptive statistics for the site location study of German Mesolithic sites

Variable	*Mesolithic Sites* ($n = 170$)	*Background Environment* ($n = 3201$)	*t-statistic*
	Mean	Mean	
Elevation (m)	478	423	11.50
Slope (percent)	6.9	9.0	-4.90
Aspect (rescaled; degrees)	95	95	0.10
Local relief (m)	99	92	2.49
Angle of view (degrees)	244	186	10.51
Shelter index	2735	2256	11.83
Horizontal distance to nearest water (m)	272	199	6.09
Vertical distance to nearest water (m)	23	13	5.48
Horizontal distance to nearest third-order stream (m)	757	709	1.21

elevations, suggesting high-elevation ridge crests and the edges of plateau tops as the primary locus of site placement in the region. Although there was no strong preference for orientation or aspect, the remaining variables were supportive of the suggested pattern. The sites possessed wider views and lower values for shelter (reflected by a higher index in Table 7.3) than the background environment, which is consistent with these high-point locations. Moreover, the results showed fairly strong tendencies for site location relatively far from water, also pointing to ridges and plateau edges, which tend to be located far from water.

A multivariate model of the Mesolithic site-locational pattern was developed during this study, not for prediction purposes but in order to assess the locational pattern in the known site sample further. A robust nonparametric discriminant function known as logistic regression (see Chapters 5 and 8) was used to develop the model, which supported the univariate findings. The model, in conjunction with the GIS, was used to map the quantitative environmental pattern of site location over the remainder of the study region (i.e., every 100 m) in order to provide a visual representation that summarizes the Mesolithic tendency (Figure 7.4b; see Chapter 10 for a more detailed discussion of how this is accomplished). The mapped pattern also supported the univariate findings of a tendency for sites to be located on ridge tops and the edges of plateau tops, considerable distances from drainages (compare Figures 7.4a and 7.4b).

A number of cautious interpretations can be drawn from these empirical data (Kvamme and Jochim 1988). Patterns in this nonrandom sample of sites might reflect Mesolithic locational preferences, modern collector biases, geological or other processes, or a combination of these factors. Geological processes might have introduced bias to the sample in a number of ways. Although the general patterns of landform and drainage in the study area have not changed since the Mesolithic,

alluvial deposition has occurred. If there are deeply buried sites in these areas of deposition, the sample will be biased away from locations in valley floors. Erosion, on the other hand, might have destroyed sites on steep slopes or along streams where meandering has occurred, thus biasing the sample away from steep slopes and drainage locations. Another factor influencing site visibility is modern land use. Materials in plowed fields tend to have higher visibility than those in forested areas, which biases the sample toward areas under cultivation, such as river terraces, gentle slopes, and ridge and plateau tops.

Geologic processes and modern land-use patterns have biased the efforts of modern collectors away from steep slopes and marshy valley bottoms and toward areas under cultivation or river terraces, gentle slopes, and ridge and plateau tops, and this is indeed a pattern similar to that demonstrated by the site sample (Figure 7.4b). The sites, however, exhibit a more restricted pattern in that they tend not to occur on river terraces or hill flanks, and they are found mainly on the edges of plateaus rather than on all portions of plateaus. Because the site distribution is more restricted than the pattern of areas inspected by the amateur collectors who reported the sites, Jochim and I have suggested that the observed distribution of sites appears to be partially the result of Mesolithic locational preferences (Kvamme and Jochim 1988).

Interpretations of these patterns should also take into account the nature of the archaeological sample. The sample used in this study included all Mesolithic sites recorded in the region regardless of function or season of occupation (factors that were unknown). Different site types could, of course, have varied locational requirements. As has been noted,

> The locational pattern of such a mixed group of sites is difficult to interpret. In part it represents a blending of characteristics specific to each site type and season, weighted according to their proportional representation in the sample. Since the site types and their proportions are not currently known, it is not possible to separate those different specific patterns. In this study, for example, sites showed no tendency to face any direction. It may be, however, that winter residential camps showed a tendency to face south, while sites of other seasons and functions had other characteristic orientations. The mixed sample would obscure these separate patterns [Kvamme and Jochim 1988].

Based on the results of our research, however, we concluded that the overall pattern reflects environmental characteristics common to all sites and all sets of activities and that interpretation should emphasize general advantages of such locations rather than those relevant only to certain seasons or specific activities. In the region of study these advantages may have included *(a)* wide views allowing easy spotting of game and strangers in any season; *(b)* strong breezes providing comfort in summer, reducing snow cover in winter, and helping to keep away insects; *(c)* good drainage in every season; and *(d)* light forests adapted to these exposed, dry situations, which may have offered ease of travel, hunting, and burning. Large distances from water may reflect an avoidance of riverine forests, the unimportance of riverine resources, or a major importance of high elevations. The

tendency for level ground probably represents the preference for performing activities on level ground (Kvamme and Jochim 1988).

In terms of the present volume, the multivariate model of the Mesolithic site pattern and its mapping (Figure 7.4b) can be viewed as a "predictive model" for Mesolithic sites based on existing data. The model remains untested, however, and its performance as a predictive tool cannot be evaluated until the model is applied to a sufficiently large, independent, and representative sample of Mesolithic sites from within the study region. At this point there is simply no way to determine whether the known site sample upon which the model is based is strongly biased (e.g., as a result of the unsystematic way that amateurs find sites or of geological processes), or indeed whether it is representarive of the region's Mesolithic pattern in general. Before the adequacy of the model could be assessed, some form of random sample survey would have to be conducted within the region, and a sufficiently large sample of Mesolithic sites would have to be discovered. The multivariate model of site location could then be applied to this new and representative sample, and the percentage of correctly predicted sites could be determined, along with statistical confidence limits around the prediction.

REFERENCES CITED

Binford, Lewis R.
1964 A Consideration of Archaeological Research Design. *American Antiquity* 29:425–441.

Clark, Philip J., and Francis C. Evans
1954 Distance to Nearest Neighbor as a Measure of Spatial Relationships in Populations. *Ecology* 35:445–453.

Dacey, Michael F.
1973 Statistical Tests of Spatial Association in the Locations of Tool Types. *American Antiquity* 38:320–328.

Dixon, W. J., M. B. Brown, L. Engelman, J. W. Frane, M. A. Hill, and R. J. Jennrich
1983 *BMDP Statistical Software*. University of California Press, Berkeley.

Hord, R. Michael, and William Brooner
1976 Land-Use Map Accuracy Criteria. *Photogrammetric Engineering and Remote Sensing* 42:671–677.

Kvamme, Kenneth L.
1986 The Use of Geographic Information Systems for Modeling Archaeological Site Distributions. In *Geographic Information Systems in Government*, vol. 1, edited by Bruce K. Opitz, pp. 345–362. A. Deepak, Hampton, Virginia.

Kvamme, Kenneth L., and Michael A. Jochim
1988 The Environmental Basis of Mesolithic Settlement. In *The Mesolithic in Europe: Papers presented at the Third International Symposium*, edited by Clive Bonsall. John Donald, Edinburgh, Scotland, in press.

Larralde, Signa L., and Susan M. Chandler
1981 *Archaeological Inventory in the Seep Ridge Cultural Survey Tract, Uinta County, Northwestern Utah*. Bureau of Land Management Cultural Resources Series No. 11. Salt Lake City.

Lin, G. C., and T. C. Minter
1976 Bayes Estimation of Parameters of the Single-Class Classifier. In *Symposium in Machine Processing of Remotely Sensed Data*, pp. 22–28. The Institute of Electrical Engineers, Purdue University, South Bend.

Mosteller, Frederick, and John W. Tukey
1977 *Data Analysis and Regression.* Addison-Wesley, Reading, Massachusetts.

Plog, Stephen, Fred Plog, and Walter Wait
1978 Decision Making in Modern Surveys. In *Advances in Archaeological Method and Theory*, vol. 1, edited by Michael B. Schiffer, pp. 383–421. Academic Press, New York.

SAS Institute, Inc.
1985 *SAS User's Guide: Basics*, version 5 edition. Cary, North Carolina.

Schiffer, Michael B., and George J. Gumerman
1977 Acquisition of Survey Data. In *Conservation Archaeology*, edited by Michael B. Schiffer and George J. Gumerman, pp. 183–190. Academic Press, New York.

Schiffer, Michael B., and Susan J. Wells
1982 Archaeological Surveys: Past and Future. In *Hohokam and Patayan Prehistory of Southwestern Arizona*, edited by Randall H. McGuire and Michael B. Schiffer, pp. 345–384. Academic Press, New York.

Schiffer, Michael B., Alan P. Sullivan, and Timothy C. Klinger
1978 The Design of Archaeological Surveys. *World Archaeology* 10:1–29.

SPSS Inc.
1983 *SPSS-X User's Guide.* McGraw-Hill, New York.

Stone, David F.
1984 *Spatial Analysis and Prediction of Prehistoric Sites in Santa Barbara County, California.* Unpublished M.A. thesis, Department of Anthropology, University of California, Santa Barbara.

Thomas, David H.
1971 *Prehistoric Subsistence-Settlement Patterns of the Reese River Valley, Central Nevada.* Ph.D. dissertation, University Microfilms, Ann Arbor.

Thomas, David H., and Robert L. Bettinger
1976 *Prehistoric Piñon Ecotone Settlements of the Upper Reese River Valley.* Anthropological Papers of the American Museum of Natural History 53(3). New York.

Wandsnider, LuAnn, and James I. Ebert
1984 Accuracy in Archaeological Surface Survey in the Seedskadee Area, Southwestern Wyoming. *Halikia'i: Contributions to Anthropology* 3:9–21. University of New Mexico, Albuquerque.

Chapter 8

DEVELOPMENT AND TESTING OF QUANTITATIVE MODELS

Kenneth L. Kvamme

This chapter is about the application of methods of empirical analysis—mathematics, statistics, and computer-processing techniques—to the development and testing of models of archaeological distributions that have a predictive capacity. This chapter is written primarily for the archaeologist with a background in quantitative methods of data analysis who is contemplating the development and testing of archaeological locational models. In order to appeal to a broader base of readers, the number of mathematical equations has been kept to a minimum, extensive descriptions of the various methods have been provided, and figures have been used to illustrate the techniques whenever possible.

Past peoples left behind material evidence of their actions—the archaeological record. This record is full of telltale patterns. Today we have access to a host of advanced tools for analyzing such empirical patterns: the tools of multivariate data analysis and the great analytical engine, the computer. We might hope to make some sense of the past by noting relationships within and among these data patterns. Using these tools I will describe in this chapter several paths toward developing and testing models of the patterns of prehistoric land use in a region.

It should be noted at the outset that formulation of rigorous models through a priori deduction of underlying causal processes is a laudable goal. We must temper this goal, however, with a practical outcome. The social disciplines presently lack a broad theoretical base, and therefore deductively based modeling strategies typically have little foundation. Haining (1981:88) has observed in geography, for example, that

> most geographers have had a preference for data analysis rather than rigorous model formation through prior specification of the underlying process. In Britain this tendency parallels the growing interest in problems of regional forecasting. The emergence of this interest in the 1970s is in part the result of the discipline's new quest for "relevance" at a policy level. As a research goal it elevates the methods of data analysis over those of rigorous model formulation through the need to provide answers to difficult and often inherently messy problems. Only the simplest spatial processes are capable at the present time of being given a rigorous formulation and there is a tendency for them to seem trivial and unrealistic when set against the expansive problems of predicting regional unemployment levels and forecasting the space-time evolution of epidemics.

The analogy with the archaeological problem of this volume is clear. Like geographers, archaeologists have a "messy" and expansive problem—modeling regional archaeological distributions. Like geographers, we can apply the methods of empirical data analysis to this problem of regional forecasting because these models are able to produce nontrivial results that can be used in applied, real-world contexts (e.g., Custer et al. 1986; Kvamme 1986; Kvamme and Jochim 1988; Larralde and Chandler 1981; Parker 1985; Scholtz 1981). This chapter focuses on these data-analysis modeling approaches.

The unit of analysis in this chapter is the *location*, or land parcel. Treating the land parcel as the unit of investigation allows greater freedom in the definition of the dependent variable used in analysis (Carr 1985:116). At the very simplest level, a binary dependent variable can be defined and coded according to whether an archaeological site is present or is not present in a particular parcel (and it can be left up to the researcher to define what constitutes an archaeological site). Some investigators (e.g., Dunnell and Dancey 1983) argue against use of the site concept, pointing out that the term *site* typically refers only to clusters of artifacts, a mere subset of the archaeological record. By using the land parcel as a focus the researcher can define virtually any archaeological manifestation of potential interest as the dependent variable. Other examples of dependent variable categories include parcels with 20 or more artifacts of any kind vs parcels with less than 20 artifacts, parcels with 10 or more sherds vs parcels with less than 10 sherds, or parcels with any cultural manifestation vs parcels without prehistoric evidence. Note that more than two categories also are appropriate, allowing investigation of multiple site or functional land-parcel types simultaneously (e.g., settlement, temporary camp, kill site, other archaeological evidence, no archaeological evidence). By using the land parcel we are able to examine various environmental, social, or other characteristics of the parcels that are coded as having archaeological manifestations or specific types of manifestations, as opposed to parcels that contain little or no archaeological evidence. An additional benefit of using the land parcel is that the size of the parcel controls the scale of investigations: very small parcels allow investigation of microenvironmental and other small-scale influences on archaeological distributions and potentially allow greater detail and precision in modeling; large parcels allow similar pursuits but on a grosser scale. (Note that if small parcels are used, and large archaeological sites or scatters are present, then contiguous parcels may be coded as "site" or "scatter" present.) In the following pages, discussion principally focuses on the simplest two-category situation for the dependent variable, for ease and clarity of presentation. All of the methods generally apply, of course, to situations in which any number of categories are being used. Since archaeologists traditionally have used the site concept, I use the term *site* in a general sense to refer to land parcels possessing the archaeological manifestations of interest, however defined. Similarly, the term *nonsite* is used to refer to land parcels that do not meet the definition of the archaeological manifestations.

The phrase *predictive archaeological model*, which has recently come into usage, is somewhat misleading because most data analysis approaches do not really predict

where as yet undiscovered sites are specifically located. Instead, data analysis approaches attempt to abstract the locational pattern exhibited by a sample of site-present locations (or specific site-type locations) in a region in terms of environmental, cultural, or other variables, and then to project this pattern over the entire region (using various computer mapping techniques, if available; see Chapter 10). If the initial sample of site locations from which the model is abstracted exhibits a locational pattern similar to that of the remainder of the region's sites (i.e., if the sample is a representative or random sample), and the sites are strongly patterned, then the mapping of the model can provide a very good indication of where sites will be found in the rest of the region. Thus, we do not predict the locations of undiscovered sites; we merely map locations that possess environmental or other characteristics that are similar to those of the initial site sample.

The nature of this mapping or extrapolation of an archaeological locational pattern depends primarily on the quality and type of modeling approach used. The mapping might correspond with simple environmental categories, such as plant communities (Figure 8.1a), or it might plot a complex multivariate function of a variety of factors with estimates of site sensitivity every 50 m across the region (Figure 8.1b). These products of empirical data-analysis models (Figure 8.1) should include performance indications—statistics that describe how well (e.g., how accurately) the model and resulting map portray the locations of sites.

It should be emphasized that the ability to predict locations (land parcels) where archaeological sites are likely to be located logically implies the ability to predict where sites are not likely to be found. Without this ability the modeling exercise becomes meaningless. It is easy to develop a model, for example, that predicts the locations of all sites within a region with 100 percent accuracy; such a model would simply classify every location (i.e., every land parcel) within the region as likely to contain sites. Of course, nothing is gained from such a model. The usefulness of a model must be judged not only by how well it predicts locations likely to contain sites but also by how well it predicts locations unlikely to contain sites. If a model is able to predict 90 percent of the site locations correctly in a region representing only 50 percent of the total land area (as opposed to 90 percent of the land area), then something is gained.

Many of the locational modeling approaches discussed in this chapter make use of basic pattern-recognition principles and techniques (Duda and Hart 1973). Predictive archaeological models developed within this perspective must work if two assumptions can be met. The first assumption requires that the locational patterns exhibited by the initial site (or site-type) sample used to "train" the pattern classifier (the quantitative model) are reasonably representative of the site population under study. The second assumption is that the site locations are nonrandomly distributed with respect to the environmental or social factors under investigation. Use of some form of random sampling designs (Mueller 1975) will usually ensure that the requirements of the first assumption are met. With regard to the second assumption, it is a basic premise of modern archaeology that human behavior is patterned, and the investigator's familiarity with the region or with

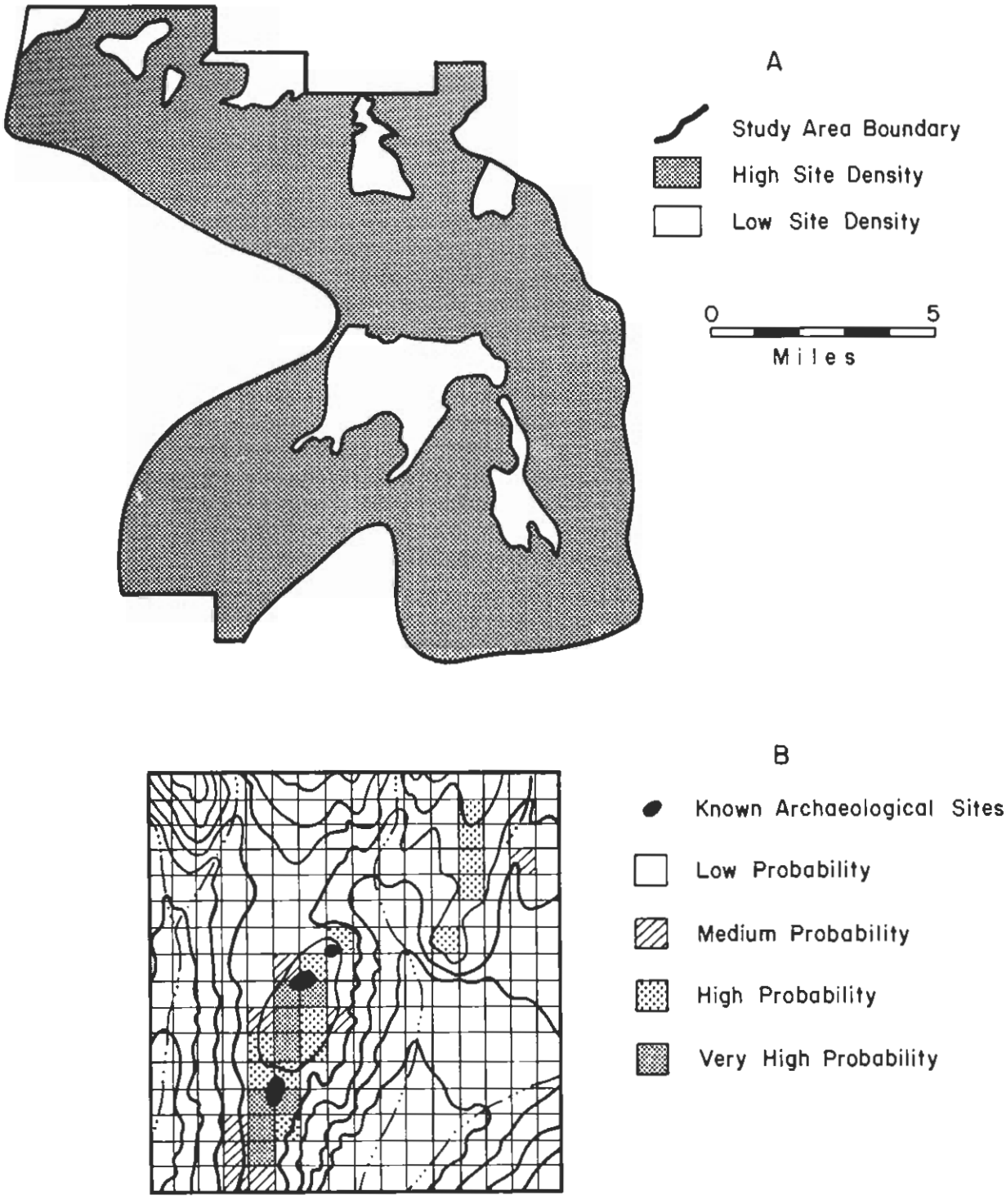


Figure 8.1. End products of cultural resource modeling. (A) A simple plant community mapping in which the communities correspond to different site densities (after Plog 1983:64). (B) A "site probability surface" superimposed on a map and derived from a complex multivariate function of six variables measured in each 50 by 50 m cell (after Kvamme 1980).

settlement data in general will usually guarantee that some of the variables selected will reflect this nonrandom behavior. When I indicate that such models "must work," I mean that there must be some *gain* (e.g., in terms of percent correct predictions) over a purely random model with no predictive capacity.

We might define the gain concept more rigorously for purposes of this chapter. It was stated above that the results of archaeological locational models should be mappable within the region under study. When the model is mapped (e.g., Figure 8.1), certain areas of the region are indicated as being more likely to contain sites than other areas. Only a percentage of all the sites (or of the site type under investigation) in the entire region will occur within the areas indicated on the map. If the area likely to contain sites is small (relative to the total area of the region) and if the sites found in that area represent a large percentage of the total sites in the region, then we have a fairly good model of site location. On the other hand, if the area predicted to contain sites is a relatively large portion of the total area and the percentage of sites within that area is not significantly greater than the percentage of regional coverage, then the model is not very useful. Based on these considerations we might explicitly define *gain* as

$$\text{Gain} = 1 - \left(\frac{\text{percentage of total area covered by model}}{\text{percentage of total sites within model area}} \right)$$

As gain approaches 1, the model has increased predictive utility; if it is near or approximately 0, then the model has little or no predictive utility. If gain is negative (<0), then the model has *reverse* predictive utility (i.e., a greater density of sites occurs outside the area specified by the model). Such a model could still be of some use if the area outside that specified by the model were subsequently considered to be the area being modeled (but the model developer should be fired!).

The gain statistic is used throughout this chapter as a means of comparing models. Most archaeological modelers tend to focus on percent correct predictions for sites, for nonsites, or even an overall percent correct statistic (see Chapter 3). These statistics can be useful and important, but they can also lead to serious misinterpretations. In addition, they offer little basis for comparisons between models (these issues are discussed in detail below), while the gain statistic presented here is easy to interpret and facilitates comparison.

An important consideration that must be addressed before model development is discussed is exactly what types of sites or archaeological manifestations are to be modeled. A central assumption in archaeology is that the locations of sites of different functional categories or chronological periods will represent responses to different situational contexts, such as environmental circumstances. It is important, therefore, to develop models for specific archaeological types whenever possible.

In practice, specific site-type models are often difficult to establish for several reasons. The problems lie not in the modeling techniques but in the definition of meaningful and justifiable site types, in assigning sites to the types based on limited and often questionable evidence, and in acquiring sufficiently large samples of the

types for subsequent analysis. The practice of assigning sites to functional types on the basis of surface information or limited excavation data is often questionable. In many regions, particularly where surface evidence consists of only a handful of lithics, the investigator may be relying on the flimsiest of evidence (if any) and on sheer guesswork. Although sites may be forced into type categories under certain circumstances, the quality of the resultant groups and their utility for subsequent analysis must be questioned. In other words, meaningless site types will yield meaningless analysis results.

A second difficulty involves categorization of sites into many site type groups, a procedure that can introduce sample-size problems. On the other hand, even when only a few site-type categories are employed, certain types within a region, such as major village centers or Paleoindian sites, might inherently exist only in small numbers. Since locational models derived from empirical data require relatively large samples in order to define a locational pattern successfully and extrapolate it to a larger region, functional or temporal types containing few cases simply cannot be modeled. In general, empirical models can be developed only for the few types that contain a significant number of representative cases. Careful thought should be given to the nature of the available evidence and the reliability of resultant site types prior to subjecting the types to a modeling exercise.

In other publications (Kvamme 1983a, 1985a) I have suggested an alternative to the practical problem of making traditional temporal-functional site types operational using regional survey data. Site types can be defined on the basis of *amount* of inferred activity occurring within a land parcel, rather than *types* of inferred activities. The amount of activity is measured in terms of quantity and variety indices of observed artifacts at a location. Locational studies can then be carried out by comparing environmental characteristics among locations indicating much prehistoric activity, locations indicating little prehistoric activity, and locations indicating no prehistoric activity. This approach allows one to investigate why certain locations were used in the past and why other locations were not.

Historical site location model development poses problems similar to those encountered in prehistoric model development, but here additional problems arise. In most regions the amount of time allotted for historical site model development is probably best spent researching historical documents and archives, which often indicate exactly where many types of historical sites are located (see Chapter 7). Moreover, the best predictor of historical site locations in many regions may be neither environmental phenomena nor the typically used cultural factors (such as distance to nearest road), but simply the cadastral survey grid, since patterns of settlement were often dictated by section and partial-section boundaries (e.g., Scholtz 1981:220). This is not to say that successful models for historical sites cannot be developed utilizing the usual environmental or other predictors. Scholtz (1981), for example, was able to construct a model for domestic historical site locations by correlating 15 environmental variables with the locations of known sites in a southern Arkansas region. Using a somewhat different approach, Monroe et al. (1980) developed powerful trend surface models for the spread of historical settlement in colonial Connecticut.

At the extreme, and depending on the quality of the regional data, models can be developed for the locations of all sites as a single group within a region. This approach has been criticized (and in certain contexts rightly so) because lumping sites of many different functional types and temporal periods into a single group introduces a great deal of variability to any analysis, making it more difficult to develop a successful cultural resource model. As we shall see in later sections, however, this variation usually is substantially less than the variation present in the environment as a whole, and it is possible on the basis of a general model to define significant portions of a region that are unlikely to contain sites of any kind. If we lump together all environmental and other variation measured at all site locations, the resultant characteristics might define an *activity space* (see Kvamme 1985a), a subset of the whole environment within which the bulk of human activity (aside from moving from one activity place to another) is performed. Although different functional activities might be conducted in entirely different situational contexts within the activity space, the activity space can be a useful construct for locational modeling purposes if it is substantially smaller than the whole environmental range of a region.

It should be recognized that the goals of cultural resource management may not always be consistent with traditional archaeological perspectives. For example, cultural resource managers are often interested in regional models for the locations of all sites in general, simply because all sites are initially important from a management standpoint. Additionally, models for traditional site types might not be as important as models for significant sites, where significance is defined as those sites being important to predefined regional research questions.

In the following pages, site location models are often referred to in a general sense. Such statements should not be taken to apply only to models for all sites as a single group, but also to models for specific types of sites, since the methods discussed are applicable to any class or classes of sites.

Finally, since this chapter covers such a wide diversity of topics, three data sets are used to provide the best possible illustrations of the methods employed. The data sets are *(a)* a western Colorado data set from a mesa and canyon region known as Glade Park, used to illustrare model-building and model-testing procedures; *(b)* an eastern Colorado plains data set, used to compare different types of modeling approaches and their mappings; and *(c)* a Mesolithic data set from the Federal Republic of Germany, used to illustrate modeling multiple archaeological sire classes.

VARIABLES USED IN LOCATIONAL RESEARCH

A researcher usually selects a variable for investigation in locational analyses because distributions of archaeological phenomena are believed to have been somehow influenced by thar variable. Hence, most researchers rely on the results of previous and similar studies in order to determine the variables to be used in an

investigation. A multitude of perspectives have been applied in archaeology to examine site locational information. Those that focus on the physical environment and its effect on settlement behavior occupy a major portion of the locational analysis literature. The examination of site catchments, topography, vegetation, and other environmental features are major elements of this approach. Roper (1979a) has labeled analyses in this perspective the study of *man-land* relationships, as opposed to *man-man* relationships. The latter term refers to analyses that assess the importance of the human or social environment in structuring patterns of settlement. These analyses focus on such areas as central place theory, the rank-size rule, and population distributions over the landscape. Although man-man relationships play a major role in the settlement pattern of modern industrialized society (Haggett et al. 1977) and offer an important and useful perspective in many archaeological applications (Flannery 1972; Johnson 1977), many key features of this approach are meaningless in a large number of archaeological situations. For example, in most hunter-gatherer contexts markets and central places are not meaningful concepts. Moreover, the primary orientation of man-man approaches is the analysis of properties related to *fixed* settlements in space, again precluding investigation of much of prehistory (e.g., many hunter-gatherer groups). In contrast, man-land relationships are intimately related to site location decisions among hunter-gatherer groups (Bettinger 1980; Jochim 1976; Wood 1978), and they play a significant role in the settlement patterns of more complex societies (Green 1973; Grossman 1977; Hill 1971; Hudson 1969). An investigation of man-land relationships can contribute to our understanding of locational behavior regardless of cultural form, and this is why most work in site locational modeling has focused on environmental data. Another reason for this focus is that environmental data are generally easier to acquire than social data. Although social factors undoubtedly influence settlement decisions in most cultural contexts, given the nature of the archaeological record it is generally impossible in any but the best understood and preserved archaeological regions to reconstruct contemporaneity between sites, population structures, etc.—important requisites for investigating social phenomena. For this reason, social factors often cannot be examined as frequently as environmental factors in archaeological locational studies.

Archaeologists have traditionally relied to an extraordinary degree on the use of nominal-level variables to describe phenomena under investigation in regional research. Examples include a focus on biotic communities, soil classes, or the practice of classifying a region as "level" or "steep." Landforms are often categorized into discrete types, such as riverine, arable, mesa top, mesa side, mesa bottom, and southern aspect (e.g., Euler and Gumerman 1978; Gumerman 1971; Plog 1971; Plog and Hill 1971; Zarky 1976); indeed, interval-level data are sometimes rescaled to the nominal level. Yet most archaeological phenomena are eminently quantifiable. Geographically distributed phenomena, particularly characteristics of the natural environment, by their very nature are distributed in a continuous manner (and thus are potentially quantifiable). Slope, aspect, and distance to nearest drainage, for example, change continuously as one moves over the landscape. Likewise, so does vegetation diversity, density, and biomass, as well as soil pH and mean grain

size. The use of categorical data and the practice of rescaling interval-level measurements to the nominal level causes critical information to be discarded, reduces the power of subsequent analyses (since nominal-level data contain less information than corresponding interval-level data), and precludes use of many powerful analytical alternatives and research designs.

A major focus of this chapter will be on the use of continuously measured data in site location research, and emphasis is placed on the importance of developing suitable measurement concepts. The types of phenomena typically investigated in this research might conveniently be grouped according to two major classes (see Plog 1971:47–48): *environmental* factors and *social* factors. The following discussion of a number of key variables that have been frequently examined in site location studies is by no means an exhaustive summary. In any particular region, some of the variables mentioned may not be appropriate.

Environmental Factors

Landform and landform-related phenomena are commonly considered in archaeological studies. A typical approach is to categorize the landscape into a series of nominal-level types, such as canyon, canyon floor, canyon side, cliff, mesa, plain, and slope (e.g., Vivian et al. 1980) and to observe the distribution of archaeological sites across these categories. Such categorization of continuous landscape forms, in addition to the problems outlined above, leads to problems of definition and tends to imply a definiteness about these categories that may not be warranted (Robinove 1981:240)—for example, how does one consistently delineate boundaries around a construct such as an arroyo head? Additionally, class boundaries may be totally arbitrary; a line dividing level from steep locations depends on current definitions of what is level and what is steep.

Steepness of ground is widely investigated in settlement studies because settlements typically are located on level surfaces where steep slopes do not interfere with activities (Judge 1973:133; Roper 1979b:77–81; Williams et al. 1973:230). This concept is easily made operational as a quantitative variable in a variety of ways, such as *slope as percent grade* (Figure 8.2a; note that the U.S. Geological Survey provides a template that performs this calculation). The form or roughness of local terrain has also been investigated (Hurlbett 1977:25–26; Plog 1981:49), presumably because rough local terrain would inhibit day-to-day activities and travel to and from sites (Ericson and Goldstein 1980). One measure of local terrain roughness is termed *local relief* (Hammond 1964); it is measured as the range in elevation within a predefined radius of a location under investigation (Figure 8.2b). High values suggest rugged terrain while low values suggest gentle terrain. A *terrain texture* measure, borrowed from image processing (Moik 1980:233), provides another alternative. An elevation is estimated at the locus of interest and at a fixed pattern of points surrounding the locus (Figure 8.2c). The variance of these elevations is then computed. High values suggest variable and dissected terrain, while low values indicate a level, smooth surface.

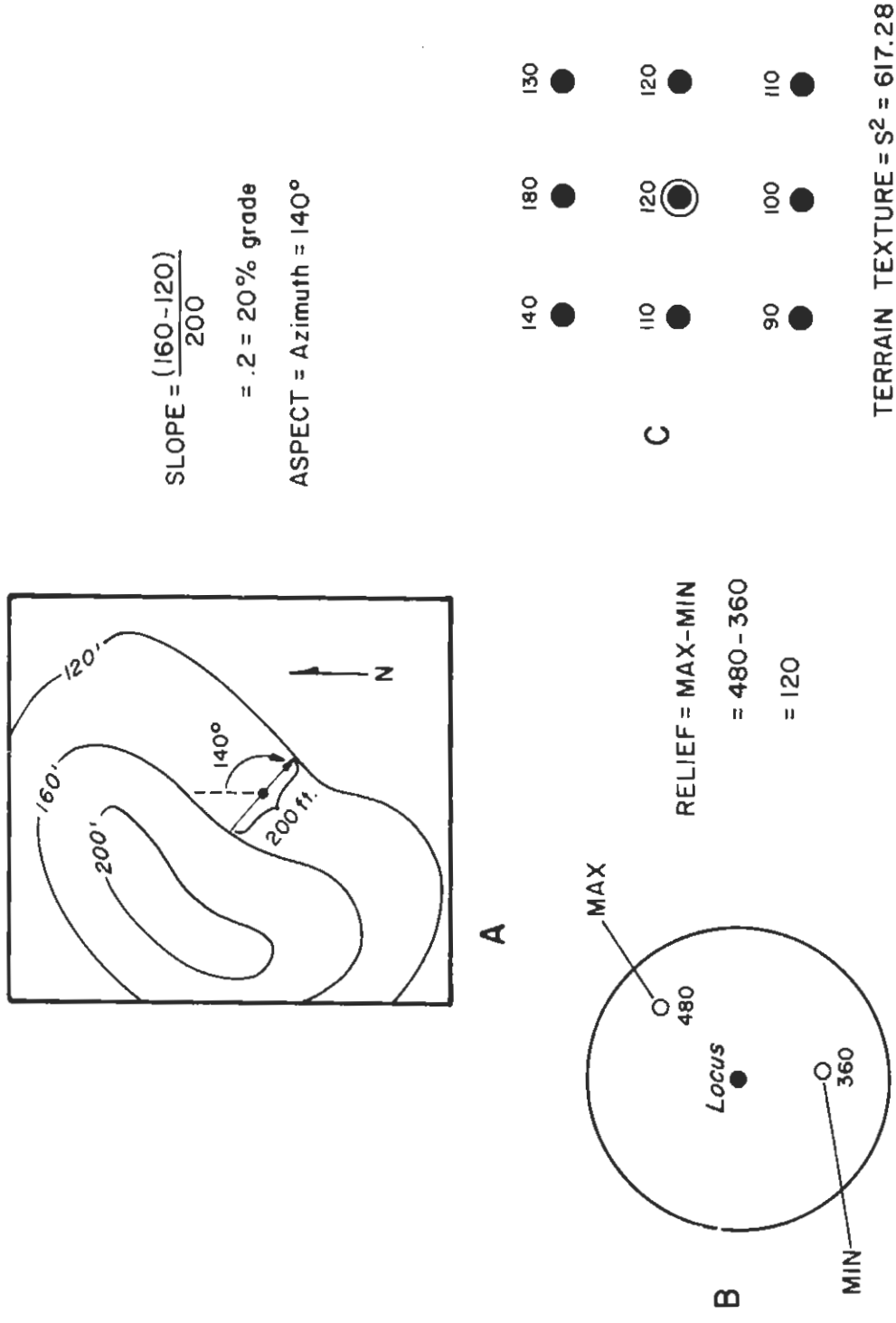


Figure 8.2. Measurement of variables: (A) slope as percent grade and direction of maximum slope (aspect); (B) local relief, the maximum elevation change within a fixed radius of a locus of interest; (C) terrain texture, the variance (s^2) of nine surface elevations recorded in a fixed pattern.

Water resources are widely viewed as important factors in locational studies. Roper (1979a) states, "some resources, such as water, are so basic and so vital that the distance to obtain them must be minimized." In a cross-cultural study of criteria influencing hunter-gatherer site-placement decisions, Jochim (1976:55) designates proximity to water sources as a central factor in determining immediate site placement. Most often examined in settlement studies are distances to a variety of water source types, such as permanent rivers, seasonal streams, lakes, springs, or streams of specified rank (e.g., Brown 1979; Judge 1973:120; Lovis 1976; Parker 1985; Roper 1979b:81; Scholtz 1981). *Linear distances* are easy to measure; *least-effort travel distances* are somewhat harder to estimate (Ericson and Goldstein 1980). Archaeologists using categorical variables generally assign class boundaries to drainage basins and let the highest stream rank in each basin represent the class category (e.g., Plog and Hill 1971:23; see Unwin 1981:79–84 for a discussion of systems of stream ranking).

The importance of view to hunter-gatherers for surveillance of the surrounding terrain is a widespread notion, and the necessity of a good field of view for spotting game animals is often cited. Jochim (1976:51, 55) suggests that a good view is one of the chief noneconomic objectives in the selection of immediate site locations among hunter-gatherers. In a more complex social context, among the pastoral Maasai (Western and Dunne 1979) view is mentioned as an important settlement location criterion purely for aesthetic reasons. A good view might also be of importance for social or defensive reasons.

A measure of *view quality* was introduced by Brown (1979:197) in a study of settlement patterns in western Kansas. This measure, which yields an angle "of surrounding terrain visible from a site" (Figure 8.3a), has been used in a number of archaeological studies (e.g., Kvamme 1983b, 1983c, 1984; Larralde and Chandler 1981; Reed and Chandler 1984). A more common measure pertaining to the view concept is a *linear distance to an overview or vantage point* (e.g., Brown 1979:197; Judge 1973:133; Larralde and Chandler 1981:118), where vantages are defined as high points, such as hilltops, ridge crests, or mesa and canyon rims. If view was important to the prehistoric occupants, then sites might be located on or in proximity to vantages. The importance of view, of course, might vary with cultural type, site function, the kind of animal being hunted, and from region to region and season to season.

Shelter and the quality of the shelter provided by a location is often recognized as being important in site location studies. Locations offering protection from wind, adverse weather, or even sunshine (in desert regions) might have been sought after for site placement. Euler and Chandler (1978), for example, examined the shelter quality of settlements in the Grand Canyon in Arizona. Among hunting-gathering groups, Jochim (1976:51) designates shelter as a central factor in the choice of location.

Shelter is a difficult concept to make operational; Euler and Chandler (1978) examined situational categories of shelter in the Grand Canyon, and Larralde and Chandler (1981) used an ordinal rank of 11 sheltering categories (from low or no

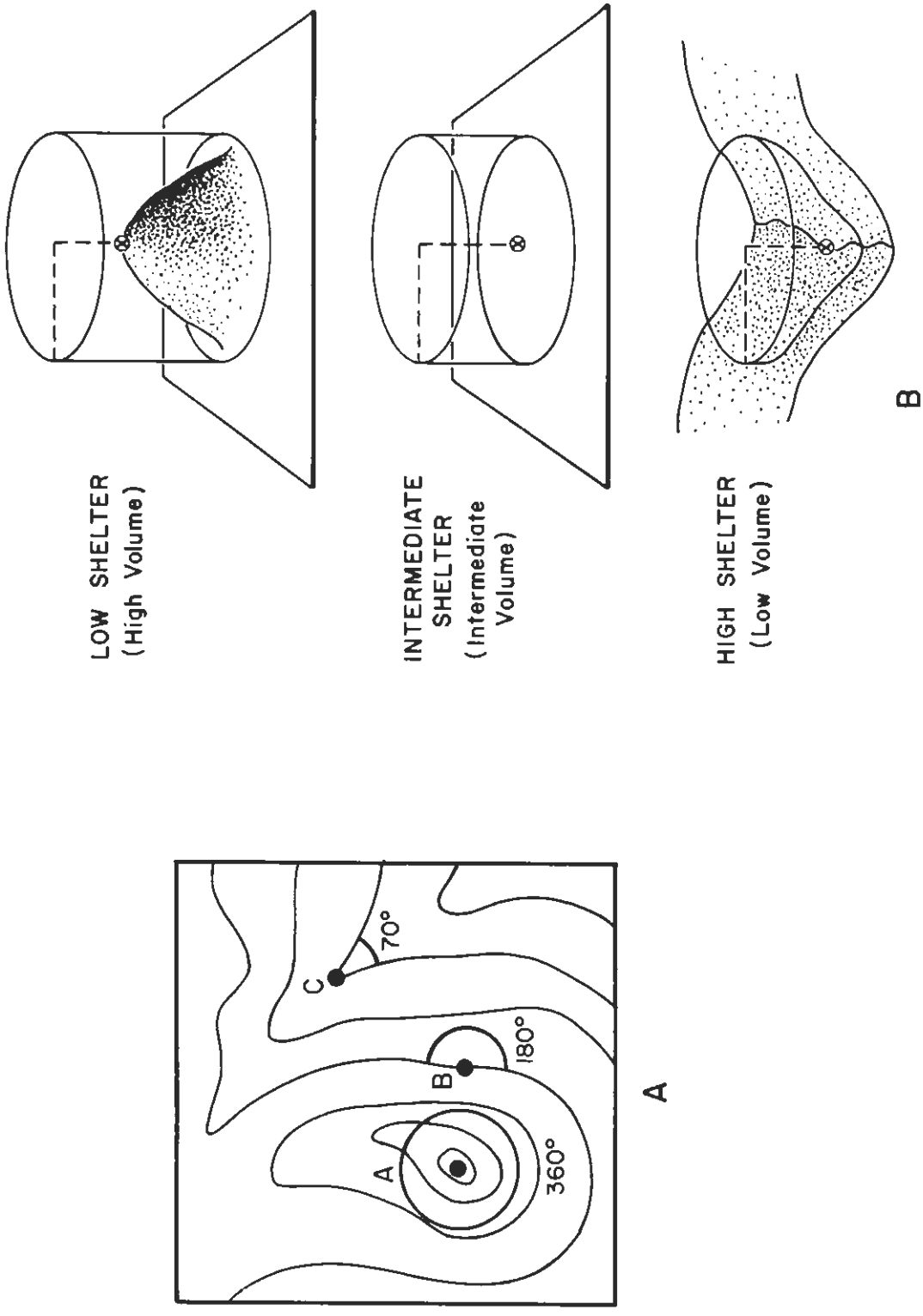


Figure 8.3. Measurement of variables. (A) View angle, a measure of view quality. The hill (A) has the widest horizontal view, the ridge flank (B) has the narrower view, and the drainage (C) has the narrowest view. (B) Cylinder volumes, inversely proportional to the sheltering effects offered by a location. The hilltop (top) offers poor shelter and has a large volume; the level plain (middle) offers intermediate shelter and has an average volume; and the valley bottom is more sheltered and has a small volume.

shelter to extremely high shelter) for site location investigations in Utah. In a recent paper (Kvamme 1984) I have attempted to devise an *interval-level measure of shelter* by considering how exposed a location is in terms of the shape of surrounding terrain. The measure is derived by imposing an imaginary cylinder over the location of interest. The top of this cylinder is a constant height (x) above the locus, and its sides are a constant distance (y) from the locus. The volume of air above the ground surface encompassed by this cylinder constitutes the measure of shelter. A large volume (e.g., surrounding a hilltop location) suggests an exposed location with a low level of shelter, and a small volume (e.g., surrounding a valley bottom location) suggests a relatively sheltered location (Figure 8.3b). The ground surface is roughly approximated by nine elevations measured at a locus of interest (0) and at surrounding loci every 45° at a fixed radius (y). The area of the base can be approximated (base = πy^2) or calculated exactly (base = $[\sqrt{8}]y^2$). The volume within the imaginary cylinder above the ground surface is calculated (after simplification) as follows:

$$\text{volume} = (\text{base}/12)(12x + 8[E0] - E1 - E2 - E3 - E4 - E5 - E6 - E7 - E8)$$

where E0, E1, etc., are the nine elevations. This index might be referred to more appropriately as an index that reflects hill-like vs valleylike characteristics (see Kvamme and Jochim 1988).

The exposure or *aspect* of a site is often examined in site location studies in connection with sheltering effects. A south-facing aspect, for example, tends to offer greater warmth from the sun (during much of the year in most of the northern hemisphere). Grady (1980:170) argues that sites may be located with primary exposures away from prevailing wind or storm approaches.

Aspect is usually measured by drawing a line perpendicular to the elevation contours of sloping terrain and recording the azimuth of this line, which provides a measurement that ranges from 1 to 360° (Figure 8.2a). A difficulty that this scale poses is that 1° and 359° both indicate approximate north, yet in a quantitative analysis 359 is much greater than 1. This difficulty can be resolved by collapsing the west half of the compass scale over the east half, such that every azimuth on the west half is given the azimuth of its mirror image on the east half. This transformation allows the measurement of direction relative to north or south where 0° is north, 180° is south, and 180° is twice as far south as 90° (east or west). Another approach is simply to use the cosine of the angle of prominent direction (Hartung and Lloyd 1969).

Resources (other than water) and their importance to site placement are often examined in site location studies. The resources usually investigated are biotic communities. A major approach is to divide a study region into *environmental categories*, such as plant communities, and to examine the number or density of sites in each community (e.g., Bettinger 1977; Thompson 1978). Catchment analyses utilize a variety of different perspectives. The percentage of various resource communities found within a fixed distance of a site might be examined (Findlow and Ericson 1980), or perhaps the variability of resources or indices of caloric potential of the area within that catchment might be calculated. Simple distance measures to

various resources are often utilized. Lipe and Matson (1971:134) mention that sites might "be located so as to maximize access to several resource zones"; Gumerman and Johnson (1971) investigate the biological transition zones between major communities, or *ecotones*, arguing that these zones "are also cultural transition zones." Simple distance measures to these resource zones might be utilized, such as a distance to the nearest ecotone or to a specific plant community (e.g., Bradley et al. 1984:75). Carr (1985:123) discusses other distance measures. When using biotic variables, the researcher should keep in mind that present-day vegetation may not necessarily correspond to past situations owing to changes in climate or land-use practices.

Finally, it should be recognized that other resources, such as fuel (Jochim 1976:51), might be important considerations in site location research. In the same vein, such resources as lithic raw materials (Johnson 1977:484) might exert a "pull" on settlement location, and a corresponding variable, such as distance from a lithic quarry, might be used in archaeological locational studies.

Social Factors

The variety of social variables utilized in archaeological locational studies is certainly smaller than the range of environmental factors that have been investigated. General concepts that have been examined relate to local *site densities*, *site proximities*, and *spacing*. Plog (1971:47–48) mentions the importance of density—the distance to other sites or sites of specific type—as well as distance to great kivas and other ceremonial sites in a southwestern archaeology context. The Southwestern Anthropological Research Group (SARG) computer system incorporates such social locational variables as number of sites within 1 km and number of habitation sites within 1 km of the site being recorded (Plog 1981:54). Horizontal distance to first- through fifth-nearest contemporary habitation sites was investigated by Adams (1974) in a locational analysis of Pueblo sites in southern Colorado.

Gravity models are often used in environmental analyses because settlement locations "appear to be related to movement-minimizing behavior" (Johnson 1977:489), which helps to justify arguments about locational proximity to critical resources (e.g., Jochim 1976). The same perspective can be applied to cultural features. Thus, distance to the nearest road or road intersection might be a useful variable if prehistoric road networks were culturally important, if they can be traced across a region, and if contemporaneity of sites and roads can be established. An implicit basis of central-place theory is that central places can be viewed as resource centers. Hodder and Orton (1976:108) illustrate empirical data that show decreasing site frequency with distance from a resource center.

Spacing between settlements is also a concern. Hill (1971:56) mentions "spacing due to competition with other groups for critical resources," which might fit in with certain territoriality concepts (Bettinger 1980:225; Wilmsen 1973). A major concept in many settlement studies is regular spacing characterized by hexagonal

arrangements of settlement around major centers or central places (e.g., Johnson 1972; Flannery 1972). Wobst (1976) discusses hunter-gatherer spacing requirements from the standpoint of demographic constraints on biological reproduction.

ASSESSING PATTERNS IN ARCHAEOLOGICAL LOCATIONAL DATA

Approaches to the study of archaeological site location are, of course, myriad (see Kohler and Parker 1986 for an extensive overview). Quantitative data analysis approaches might initially be lumped into two categories: those based on trends in location and those based on trends in characteristics of locations. Models of the locational trends of site distributions are based solely on spatial coordinates; locations in space are modeled, not characteristics of locations. As Parker (1985:202) notes,

> Even in the case of accurate representation of a distribution . . ., this methodology gives no information for explaining why the distribution is in a particular form. Explication of site settlement systems is enhanced by methodologies which relate site presence to location characteristics, thereby allowing interpretations as to why sites are located where they are.

Models of trends in locational characteristics, on the other hand, analyze empirical relationships among characteristics of the natural or social environment and the locations of sites. Modeling of locational characteristics has been the dominant approach, and such models are to be preferred not only because of their generally greater power (see below), but also because they offer some potential for interpretation.

Approaches Based on Trend in Location Only

Approaches that focus on trend in location attempt to model regional site distributions only on the basis of locational (x,y) coordinates. No other information is used. Positions in space are modeled, not characteristics of the spatial positions. Hence, these models are generally rather crude.

Trend-surface analysis (Unwin 1975), a regression technique, is one procedure for modeling locational trends, although it is not ideal for site location data. Based on spatial coordinates of known sites, most archaeological applications develop functions to model a *continuous* dependent variable, such as trends in dated sites, across a region (Bove 1981; Monroe et al. 1980; Roper 1976). Other examples include modeling trends in length/width indices of Bagterp spearheads across northern Europe or varying percentages of Oxford pottery across southern Britain (Hodder and Orton 1976:164–174).

Note that all of these studies utilize a continuous dependent variable, which poses something of a problem for site location analysts because ofren their goal is to develop models for discrete classes of such information as site (or site-type) presence or absence. This amounts to a nominal-level dependent variable, for which most regression techniques are poorly suited. One analytical alternarive for site location modeling in the traditional regression context is to convert the presence/absence criterion to some numeric form that the technique is better able to handle. This might be accomplished by placing an arbitrary grid over the region and estimating site density or performing a simple site count in each grid cell to provide a dependent variable that is more than dichotomous. This approach has been used in a number of archaeological studies to develop regression models of artifact counts per grid unit for intrasite distributional analyses (Feder 1979; Hietala and Larson 1979; Larson 1975). For site location studies, a similar approach could be applied on a larger scale by gridding a region and treating sites as the unit of analysis.

A major problem with the trend-surface regression approach is that different results can be obtained depending on which arbitrary grid size is chosen. A second problem involves the deficiencies of the regression model when it is applied to a dependent variable consisting of counts. Hodder and Orton (1976) and Davis (1973) discuss general problems in the use of trend-surface analysis.

Kriging approaches to the same problem (Parker 1985:202-205; Zubrow and Harbaugh 1978; Chapter 2, this volume) utilize similar kinds of data, spatial coordinates and site counts per gridded unit area, and generally do a better job of modeling densities across a region than trend-surface approaches (Delfiner and Delhomme 1975). This method also suffers from problems resulting from arbitrary grid sizes, however.

Recently, an approach to trend mapping that is specifically designed for nominal-level class categories has been developed (Wrigley 1977a, 1977b). This method is based on a logistic regression technique (see below) and can be referred to as *logistic trend-surface analysis.* It makes no assumptions about distributional form, and it is appropriate for a nominal-level dependent variable. Moreover, the dependent variable can consist of multiple class categories (e.g, site absent, site type A present, site type B present, site type C present). For a given locality, with spatial coordinates x and y, the outcome is a value for a class that is constrained between 0 and 1. This value can appropriately be interpreted as the probability of an outcome, such as site presence, given its location coordinates (Wrigley 1977b:12). Examples of this technique all come from geography and include the probability of households in a neighborhood shopping at a particular market vs the probability of the households not shopping at that market (a two-class problem; Wrigley 1977b) and probability trend surfaces of households highly annoyed, moderately annoyed, and little annoyed by aircraft noise in the vicinity of Manchester Airport (a three-class problem; Wrigley 1977a).

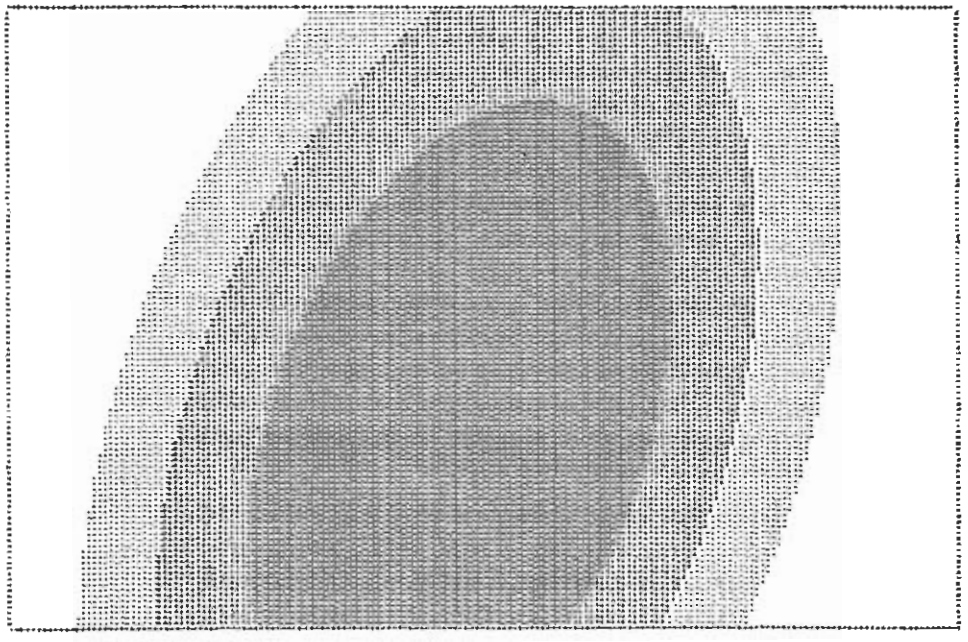
A model of archaeological site trend in location can be developed through application of the logistic trend-surface technique. The locations of 95 known open-air lithic scatters in a 5.5 by 8.5 km test study region on the southern Colorado

plains are presenred in Figure 8.4a (this study region will be extensively used for examples in later sections of this chapter). The study region has been gridded into approximately 19,000 cells (land parcels) measuring 50 m on a side; Figure 8.4a illustrates those cells with open-air sites present. The results of various efforts to develop a probability trend surface for the presence of this site type based only on the spatial coordinates of the known sites are shown in Figure 8.4b-d. This is a simple two-class problem of site presence and site absence, although we are interested only in the mapping for the site-present class. (Note that in a two-class problem the mapping of one class is the "negative image" of the other class since probabilities at any locus must sum to unity. Thus, it is not necessary to produce probability surface maps for both classes. In a problem context involving three or more classes, however, a separate probability surface map for each class is required, each derived from a separate equation that is mathematically calibrated to the other class equations.) The site-absent locations were obtained at 54 locations (cells) systematically placed every kilometer across the study area.

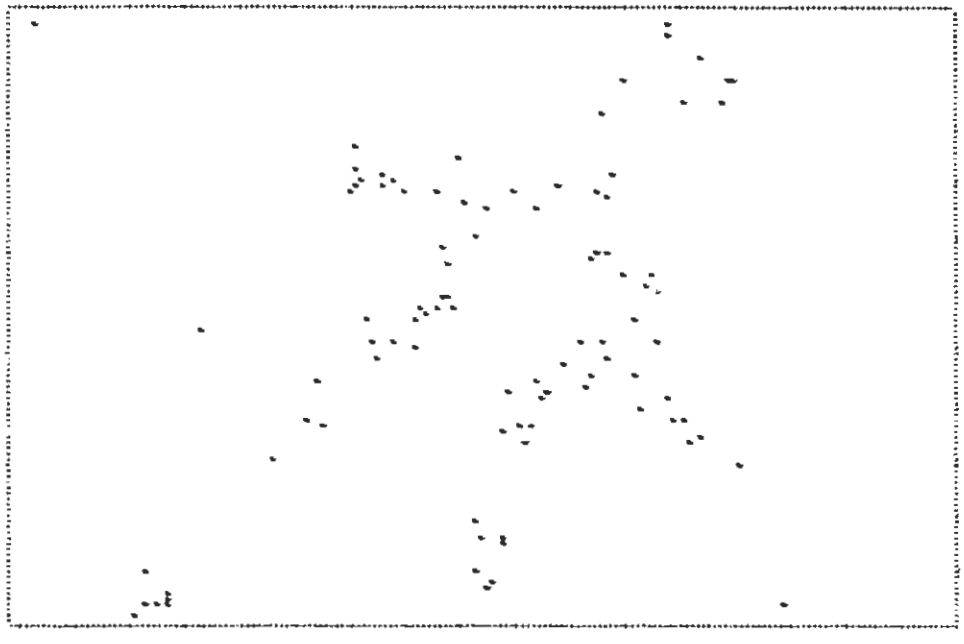
First- through fourth-order logistic trend surfaces were fitted to these data using the BMDP logistic regression program (Dixon et al. 1983). Fitting trend surfaces to empirical data requires use of polynomial functions, which employ various powers of a variable. A function of x and x^2 (a second-order model) makes a graph with one "bend"; a function of x, x^2, and x^3 (a third-order model) makes a graph with two bends, and so on. Since we are working in a two-dimensional space with (x,y) coordinates, we need to express powers of both variables ($x, x^2, x^3, \ldots y, y^2, y^3, \ldots$) plus all interactions between the two variables (xy, x^2y, xy^2, x^3y, . . .). Generally, the higher the order of the model the better the fit to the data. Because the resultant functions are only combinations of these rather meaningless variables and their powers, it becomes clear what Parker (1985:202) was alluding to in the quotation given above, when she claimed that these models have little explanatory potential.

The first-order probability surface contains the terms x and y. The second-order model adds the terms x^2, xy, and y^2; the third-order model adds to these the terms x^3, x^2y, and y^3; the fourth-order model adds the terms x^4, x^3y, x^2y^2, xy^3, and y^4 (see Feder 1979:96). Thus, the fourth-order model contains a total of 15 parameters (including an intercept) that must be estimated. Second- through fourth-order surfaces are portrayed in Figure 8.4, with site-presence probabilities portrayed in steps of 0.2 probability and in levels of increasing darkness.

In traditional trend-surface analysis (discussed above) the utility of the various polynomial surfaces are usually evaluated on the basis of increases in R^2 (variation accounted for in the dependent variable) over previous surfaces (Unwin 1975). This is not possible with the logistic trend-surface technique since the dependent variable is categorical. A number of pseudo-R^2 statistics for logistic regression have been introduced. One, R_p^2 (Baxter and Cragg 1970), provides a value that ranges between 0 and 1, although midrange values are considered very good for indices of this kind (Stopher and Meyburg 1979:334). The first- through fourth-order surfaces shown in Figure 8.4 yield the following values of R_p^2: 0.0218, 0.3125, 0.3799, and



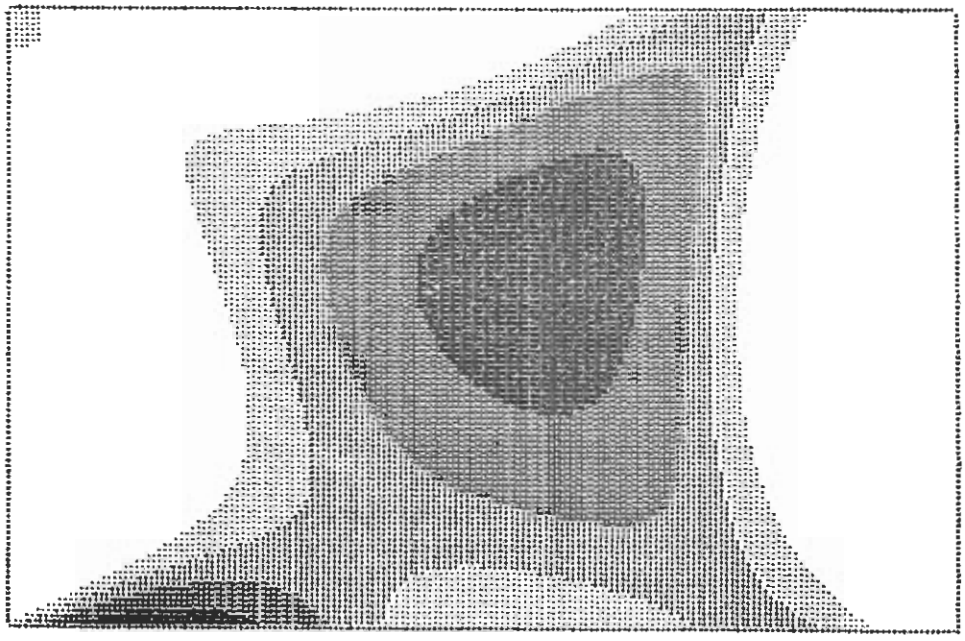
B



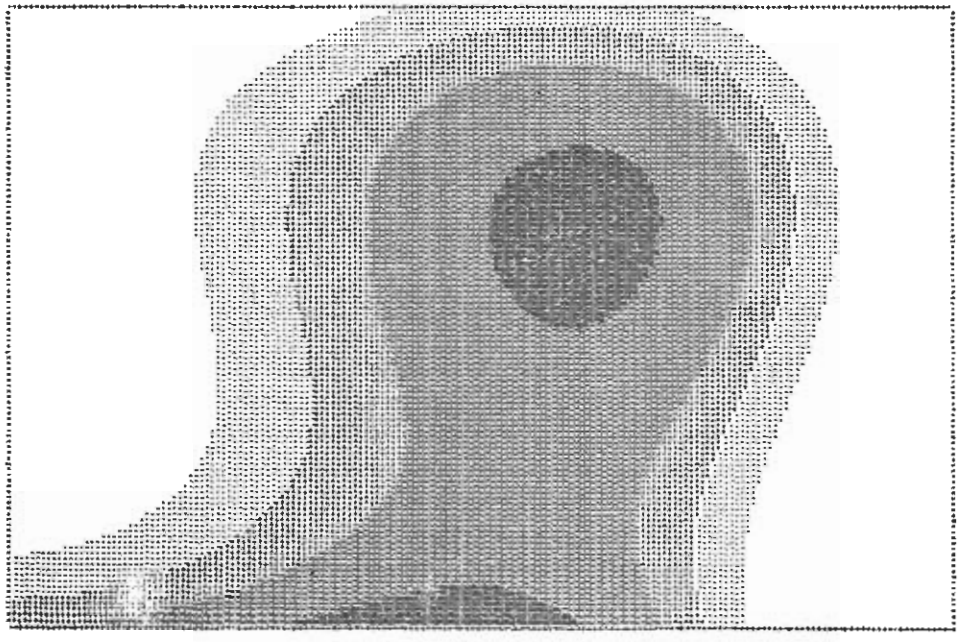
A

Figure 8.4. Models based only on locational or positional information (x,y coordinates): (A) locations of known sites, (B) second-order logistic trend surface, (C) third-order logistic trend surface, (D) fourth-order logistic trend surface.

(continued)



D



C

Figure 8.4. Continued.

0.5043, tespectively. Thus, the first-otdet surface accounts for almost none of the "variation" in site presence/absence. The second-order surface ptovides a substantial improvement (increasing R_p^2 by about 0.29) because the resulting probability ellipses center around the major concentration of site locations (Figure 8.4b). The third-order surface provides an increase in R_p^2 of about 0.07, and the fourth-order surface yields another leap, an increase of about 0.12. Note that the fourth-otder surface (Figure 8.4d) does a relatively good job of modeling or describing the spatial distribution of the known sites (several branches and clusters of sites are picked up by the surface), considering that it represents a simple function based solely on the spatial coordinates of the site-present and site-absent data. It should be apparent that if the locations of the known sites in a modeled region are representative of the locations of unknown sites in as yet unsurveyed areas of that region, then high-order logistic trend surfaces offer a predictive aspect, like any other model.

We might also apply the gain statistic, discussed above, in order to examine model performance in a more interpretable way. The gain statistic was defined as one minus the ratio of the percentage of the total area encompassed by a model when mapped, divided by the percentage of total sites within a model's area; a good model is suggested as values approach 1 (small area with a high percentage of sites). The locations in Figure 8.4 with an estimated probability of membership in the site-present class greater than 0.5 (the two and one-half darkest levels of shading) can be used as the area encompassed by each model. The 0.5 point, which is a traditional decision rule, is arbitrarily used here and elsewhere for comparative purposes only; later sections examine other decision rules. When the 0.5 level is used, the second-order model (Figure 8.4b) covers approximately 40 percent of the study region and 74 of the 95 sites (78 percent) occur within that area. This yields a gain statistic value of 1 - 40/78 or 1 - 0.513 = 0.487. A similar assessment of the third-order model (Figure 8.4c) reveals that the modeled area is 39 percent of the total area and that 80 percent of the sites (76 of 95) lie within that area. Thus, the third-order model provides only a slight improvement in gain (gain = 0.513). The fourth-order model (Figure 8.4d) provides a major improvement, encompassing only 31 percent of the total area and including 82 percent of the sites (78 of 95), yielding a gain statistic of 0.622.

Approaches Based on Trends in Locational Characteristics

Archaeologists have examined trends in archaeological site locational characteristics, particularly environmental features, for a long time. To illustrate, in a study of the Paleoindian occupation of central New Mexico Judge (1973) examined water sources, vantage points (from which game might be viewed), hunting areas, and trapping areas (locations where large animals could be driven and trapped), and their relationships with the sites in his sample. In their investigation of prehistoric Shoshonean settlement patterns in Nevada, Thomas and Bettinger (1976) examined distance to water, distance to the piñon ecotone (piñon was considered an important economic resource), elevation above the valley floor, and ground slope. The

importance of shelter, fuel (firewood), a good view (to observe game), and water to the immediate locations of hunter-gatherer sites in general were outlined by Jochim (1976:55) in a study based on ethnographic literature. Analyses pertaining to more complex agricultural situations often examine conditions related to the arability of the land. For this reason, Green (1973) investigated five variables related to soil type in an analysis of Maya settlement in Belize. A soil texture variable as well as vegetation, hydrographic, and landform variables were examined by Roper (1979b) in a study of Woodland site locations in central Illinois. In all of these studies, characteristics of site locations are the focus of interest, and as noted earlier, various social factors, such as distance to the nearest contemporary road or to a site offering services or religious or social resources, may also be considered characteristic of a location.

Approaches of the kind just described typically summarize empirical data observed or measured at known site locations through tables or various descriptive statistics. The ability to "predict" in general terms on the basis of these data patterns is implicitly or explicitly recognized. Many archaeological studies of this type have depended largely or wholly on the use of nominal-level categories for investigating site location patterning. One such predictive model developed by Bettinger (1977:220) was constructed for "predicting the distribution, function, and density of archaeological materials in the Inyo-Mono region." This model simply divided the study region into biotic communities and projected expected numbers of various site types in each community based on site density estimates obtained from sample surveys. This is the most common approach in traditional site-location investigations, and discussion of other examples (e.g., Brose 1976; Grady 1980; Reher 1977; Thompson 1978) would be redundant.

Other investigators have focused on continuous site-location information (e.g., Judge 1973; Findlow 1980; Hurlbett 1977). Such empirical data can be quite useful in formulating projections about site locations. One might show, for example, that x percent of sites occur within y distance of a drainage in a study region by obtaining measurements of distance to water from a representative sample of known sites in the region. Thomas and Bettinger (1976:362–363) go a step beyond this by fitting normal distributions to empirical data on slope, distance to water, distance to piñon ecotone, and elevation above valley floor obtained at site locations in the Reese River Valley of central Nevada. The central portions of these normal curves are taken to represent "ideal locations" for sites. Moving in either direction along the curves (e.g., to steeper ground) decreases the probability that sites will be found.

The practice of fitting theoretical distributions to data is a common one in many disciplines (e.g., Cooper and Weekes 1983:20). The above procedure of Thomas and Bettinger might seem useful for modeling site distributions, e.g., projecting site probabilities across the landscape. Such models are called *single-class classifiers* in remote-sensing applications (Lin and Minter 1976; Minter 1975) because they are used to describe the distribution of a specified class (e.g., a site-present class) using data only from that class. (Such single-class approaches do not perform

as well as approaches that utilize a second class as a control group to contrast with the group of interest; this latter approach will be described below.)

A problem with archaeological studies of the type discussed above is that they often consider nominal- or interval-level variables only singly, on a univariate level. Data often are not examined in a multivariate context, and as a result interrelationships and redundancies between variables are seldom considered. Nor are their joint effects on site location taken into account for prediction purposes, even though a cursory inspection of the literature points to the multivariate nature of the site location problem.

Control Groups

An important methodological difficulty of many archaeological site location studies is the failure to use a control group with which archaeological distributional patterning may be compared. We might imagine, for example, a newspaper report indicating that "90 percent of the inmates of Smith County jail are nonwhite minorities." Such statistics are often used in lay contexts, but a scientist seeks background control data before formulating conclusions. If a control group obtained by selecting a random sample of members of the entire population of Smith County were to indicate that 90 percent of the population are nonwhite minorities, this would suggest that the jail inmate proportions do not represent a noteworthy pattern. This example has direct bearing on archaeological site location studies because the same kind of initial argument is offered in many studies, namely that x amount of sites are located within y distance of a resource.

In many disciplines, control data sets are routinely used. Quantitative psychologists, for example, typically measure personality traits on a control group selected randomly from the population. This reference body of data is then compared with data from the group under study, e.g., suicide-prone individuals (Overall and Klett 1972:257). Geologists have compared locations exhibiting high levels of radioactive emissions with a control group of locations exhibiting low emission levels in order to develop predictive models for uranium exploration (Missalati et al. 1979). Remote-sensing scientists obtain spectral data from a variety of environmental settings in order to amass a comparative background with which spectral emissions of crop types of interest, such as wheat, can be compared (Landgrebe 1978; Swain 1978). These techniques are common in pattern-recognition studies (Duda and Hart 1973).

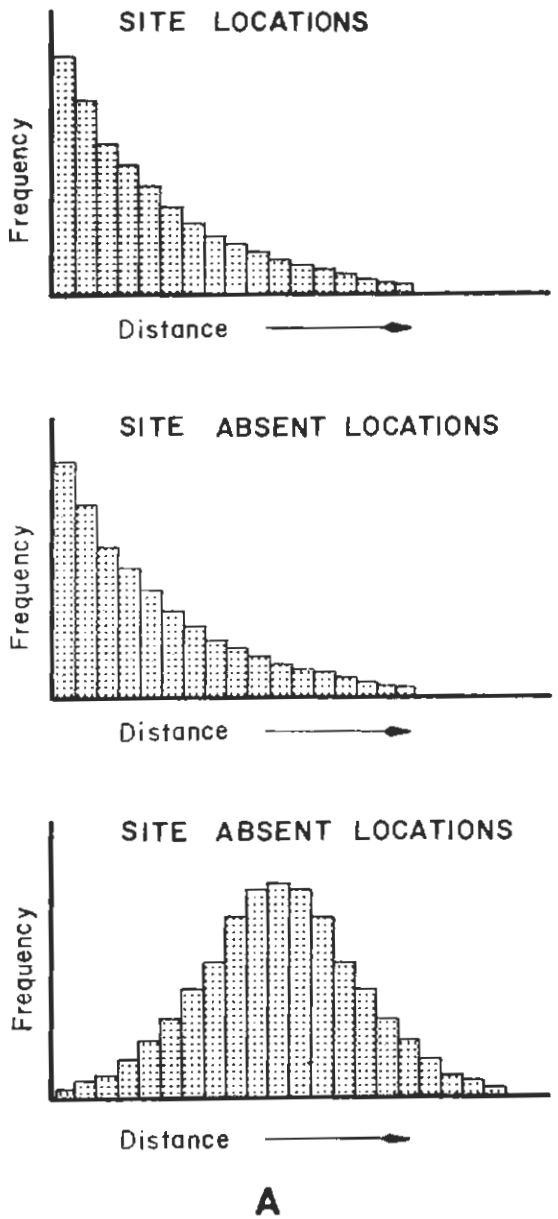
In archaeology a similar approach can be taken. Environmental or other information can be measured at locations (land parcels) containing known archaeological sites and then contrasted with a control group of identical measurements obtained at random locations in a study region where sites are known to be absent. By this means, environmental and other features beating relationships with the locations of sites might be identified. Data for such a variable as distance to nearest drainage, for example, might be collected at a representative sample of archaeological locations within a region (Figure 8.5a, top). Since the distribution of the data is

concentrated in the area of the graph representing short distances to water, a typical archaeological conclusion might be that proximity to water is an important factor in site location. Yet we must also ask how far any location within the region under study is from a water source before drawing such a conclusion. A control group of measurements of distance to water taken at random locations where sites do not occur might yield an identical distribution (Figure 8.5a, middle), forcing the conclusion that water is generally close to any location and that proximity to water is not a significant factor in site location in this area. If, on the other hand, the control data yielded a distribution with a central tendency some distance from water (8.5a, bottom), the archaeologist might more justifiably arrive at the conclusion that proximity to water is a significant locational factor (see Kvamme 1985a).

As the above example suggests, a control group approach may be essential to forming valid conclusions concerning site location factors when empirical archaeological data are used. Control groups serve several important functions. Their use in model development is discussed below, but perhaps the most important use of control group data is in model testing; it is only through the use of a control group that the performance of a site location model may be properly assessed. Returning to the example given in the introduction to this chapter, a model might classify *every* location (land parcel) within a region as site-likely and thereby predict all actual site locations with 100 percent accuracy, but such a model is useless. (In this case the gain statistic would yield 1 - [100 percent of total area classified by model]/[100 percent of sites in model area] = 0.) On the other hand, by using a control group that approximately represents the environment at large, it might be found that a site location model encompasses only 60 percent of the land area of the region when mapped, but it includes 90 percent of all sites within that area, representing an amount of gain against which the utility of the model may be judged (in this case, gain = 1 - 60/90 = 0.33; see the section below on "Assessing Model Performance" and the discussion of gross errors and wasteful errors in Chapter 3).

The use of a nonsite control group also helps to clean up some conceptual sloppiness. Through use of certain statistical procedures, we often wish to estimate the probability of site-group membership at a location (land parcel). Obviously, this probability often can be less than 1.00. But if the probability of site-class membership is estimated as 0.6, what does the remaining 0.4 probability represent? Logically, and consistently, this remaining probability represents site absence, the complement of site presence. Thus, models for site presence must also consider site absence, and nonsite data permit us to do this.

Nonsite locations should be selected from throughout the region in which the sites under question are being modeled. If the region is large and diverse, with multiple natural subgroupings of the environment (e.g., plains and mountains), then the investigator might wish to examine site location patterning within each grouping (a plains model and a mountains model). Such a practice could lead to enhanced precision of predictions. In this case, it is appropriate to randomly select nonsites according to the groupings (strata).



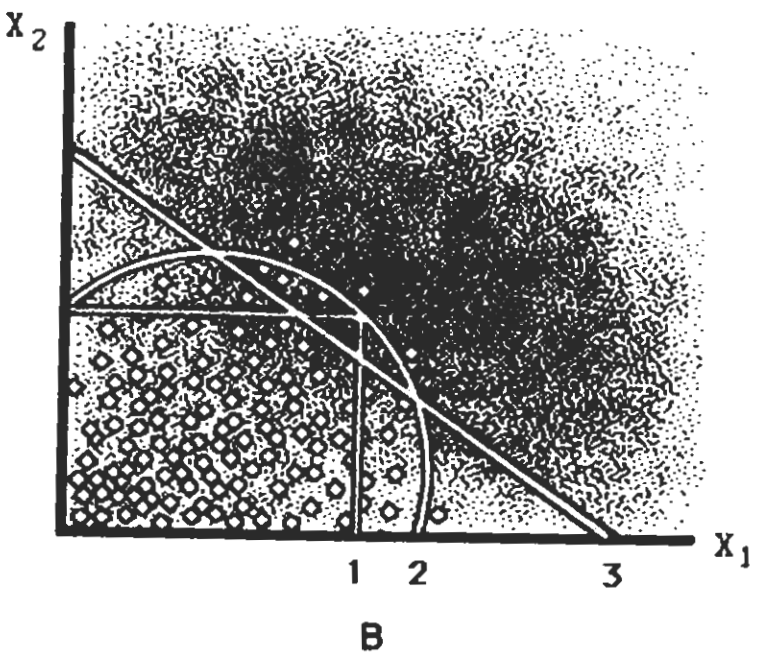


Figure 8.5. Uses of control data in empirical studies. (A) Distances to water in a hypothetical study region. The top histogram indicates the empirical distribution measured at a representative sample of site locations. If distances to water were measured at random locations where sites were known to be absent and the middle distribution resulted, it could be concluded that proximity to water is not a significant site-location factor. If the bottom distribution resulted, the importance of proximity to water would be indicated (after Kvamme 1985b). (B) A two-dimensional measurement space where X_1 might be ground slope and X_2 might be distance to water. The decision boundaries attempt to separate the site (hollow circle) and nonsite (dot) classes: (1) level-slice, (2) quadratic, and (3) linear decision boundaries (discussed in text).

The use of background data sets has been explored to some extent in archaeological site location studies. Plog and Hill (1971), Plog (1971), and Flannery (1976:92–93) point to the importance of knowing conditions in the environment as a whole before assigning significance to a particular factor in terms of site location, and Plog (1968) and Zarky (1976) actually determine background characteristics in their studies of prehistoric settlement systems. These studies focus on proportions of gross environmental categories (e.g, arable land, mesas, river bottoms) in the study region as a whole as a basis for contrast with the observed pattern of site distribution with respect to the same categories; differences in proportion are interpreted as implying some sort of selection on the part of the prehistoric inhabitants for some of the environmental categories. In contrast to this focus on large-area environmental types, which offers little information on the immediate locations of archaeological sites, a technique in which control data are measured at random "point" locations (e.g., land parcels or quadrats of very small size) at which sites are known to be absent can provide suitable background contrasts to identical information recorded at known site locations (Custer et al. 1983, 1986; Kvamme 1980, 1983b, 1984, 1985a, 1986; Larralde and Chandler 1981; Morain et al. 1981; Peebles 1981; Wells et al. 1981). A somewhat different approach, but one that uses an identical methodology, measures control data for large land parcels (e.g., one-half or one square mile) that contain no archaeological sites (Holmer 1979; Scholtz 1981; Schroedl 1984; Zier and Peebles 1982).

Patterns and Classification: The Measurement Space

Scientists working in remote sensing, pattern recognition, statistics, and decision theory have developed a number of ways to classify objects (individuals, locations) into prespecified groups. A great deal of practical experience in predictive modeling in geographic contexts has been gained by researchers attempting to analyze and classify remotely sensed images.

In image analysis studies, multispectral scanners (MSS) on platforms in orbit above the earth sense reflected radiation from the earth's surface (see Chapter 9, this volume). The predictor variables are the various MSS bands in which reflected radiation is measured. The basic unit of analysis is termed the *pixel* (picture element), which corresponds to a small area on the ground. Reflected radiation values are measured on each MSS band (the variables) for each pixel in the region of interest; image classification scientists then use the measured reflectance characteristics to classify each pixel into likely (prespecified) groups of interest, such as wheat vs nonwheat, forest vs nonforest, or urban vs nonurban areas (Landgrebe 1978).

The analogy with our archaeological problem is clear: in many site-location modeling approaches we want to classify locations (analogous to pixels and often small in area) into site-likely, site-type-likely, or site-unlikely categories on the basis of the variables (usually measuring terrain or environmental characteristics) measured at the locations. Modeling approaches that utilize computer-based geographic information system (GIS) techniques (Hasenstab 1983; Kvamme 1983b,

1986; Chapter 10, this volume) actually grid entire study regions inro small cells (pixels) and treat these cells as the units of analysis. As a resulr of this general similarity between the problems of remote-sensing classification and those of site-location modeling, many of the techniques presented in this chapter are borrowed directly from pattern-recognition and image-classification studies.

In pattern-recognition and image-analysis research, measurements obtained at locations belonging to known categories are often called *training data* because they are used to develop or "train" classification functions. These functions are numerical decision rules that utilize class characteristics (i.e., measurements) to classify entities whose group membership is not known (Swain 1978:142). In an archaeological context, samples of known archaeological site locations constitute a training set, and measurements of environmental and other variables at each of the sites provide a site class characterization. If a control group of site-absent locations is used, measurements at these locations provide a nonsite class characterization. Patterning represented by the measurements of each class can be used to assign future locations (for which site presence/absence is unknown) to one of the classes in a predictive sense. Exactly how this is accomplished depends on the nature of the technique used (several alternative methods are presented later), but all techniques for accomplishing this goal have an underlying similarity.

Archaeological locational data typically occur as a series of points or small areas on maps that represent the locations of known archaeological sites or artifact clusters. These site locations might suggest an intuitively identifiable settlement pattern; for example, the sites might be located along high terrace ridges above major drainages within stands of oak. In working with classification procedures that can be used to model a site location pattern in an objective manner, however, a more abstract concept of the term *pattern* is required. Characteristics of a location are reduced to a series of measurements (which may be categorical), and the classification procedure compares the measurements with a set of previously made measurements that are "typical" of known classes, such as site-present and site-absent categories. The location is assigned to the group whose measurements are most similar to its own. In other words, as far as a classification procedure is concerned, after the measurements are obtained the physical form of the location and of the surrounding environment are unimportant: the set of measurements *is* the environmental (or other) pattern of the location (see Swain 1978:139). In general, we might think of the environmental terrain characteristics of a location simply as a set of measurements, not in terms of their physical form.

The n measurements made at a location define a point in n-dimensional space, which is referred to as the *measurement space*. The purpose of a classification procedure is to divide the measurement space into appropriate *decision regions*, each corresponding to a specific discriminable class, and to assign the measurements made at a location to the class that corresponds to the decision region in which it falls. A two-dimensional measurement space where X_1 might be ground slope and X_2 might be distance to nearest water is illustrated in Figure 8.5b (above). The

site-present locations (hollow circles) tend to possess level ground and are close to water, while the site-absent locations (the black dots) overall tend to be on somewhat steeper ground and farther from water. The decision boundaries (several are presented for later reference) attempt to separate the two classes. If X_1 and X_2 were measured on a map at some location where site presence/absence is unknown, the location would be identified as more similar to the site or the nonsite group, depending on where its measurements fall relative to the currently defined decision boundary. Of course, some nonsite locations will always fall on the site side of the decision boundary and some sites may fall on the nonsite side, which introduces an amount of error that we attempt to minimize. (The case exemplified in Figure 8.5b is an oversimplification since, in practice, we work with many more variables [dimensions], which provide more information and help to reduce error.)

Although the above example utilizes continuous data, categorical data can be approached in the same way. When dealing with such variables the measurement space is best seen as a table, with one dimension representing class partitionings according to one variable (e.g., plant communities) and the other dimension comprising class partitionings according to another variable (e.g., topographic categories).

Practical Statistical Concerns in Model Development

In earlier chapters a great deal has been said about statistical inferential techniques and their proper application in site location studies, particularly with regard to meeting various statistical assumptions. It is often difficult, however, to meet many of these assumptions in real-world applications. This section briefly discusses certain statistical difficulties pertaining to sampling and model development.

A concern commonly voiced in regional archaeological studies pertains to apprehensions about cluster sampling and the problem of obtaining representative or unbiased samples of sites from within a region (e.g., Berry 1984; Mueller 1974; Thomas 1975; Chapter 5, this volume). What is meant by representativeness is that the characteristics of a sample (i.e., sample statistics) are unbiased estimates of the true parameters of the population under study (e.g., the mean slope value estimated from a sample of sites provides a good estimate of the true mean slope value for the population of all sites in the region of study).

Some archaeologists maintain that the only way to obtain unbiased site samples in a regional context is through simple random sampling, but to obtain a correctly drawn simple random sample of sites from within a region, the location of every site would have to be known beforehand, each would be assigned a number, and a simple random sample would be obtained by random selection of the numbered sites (see Thomas 1975:78). Clearly this approach is impractical; moreover, if every site were known prior to the sample selection, there would be no need for the site-location modeling exercise.

An alternative procedure that would allow simple random sampling requires that the researcher superimpose a small-mesh grid (where each grid cell is approximately the size of a typical archaeological site) over the region of study. Each grid cell is assigned a number, a simple random sample of cells is drawn, and this sample is then surveyed. Most of the cells will not contain sites, owing to their relative rarity (see "Base Rate Probabilities," below). If they are very rare, for example occupying only about 1 percent of the cells, then this procedure also presents difficulties since many thousands of cells would have to be surveyed to obtain reasonable site sample sizes. Additionally, the problem of traveling to and locating numerous randomly placed small cells presents a nontrivial factor that must be considered.

Even if we could obtain simple random samples in regional surveys, problems would still arise in attempts to draw statistical inferences during model development. Most techniques of statistical inference assume independent observations. Statistical independence in terms of areally distributed data implies that when observations are ordered in space it should not be possible to have a better than random chance of predicting values of some observations when other values are known. As Gould (1970:443) points out, "it is doubtful that one could find an assumption that is more at variance" with geographical data; spatially distributed phenomena generally possess regular spatial variation or positive spatial autocorrelation (Cliff and Ord 1973), thus violating independence assumptions. Tobler (1970:234) has referred to this property as "the first law of geography: everything is related to everything else, but near things are more related than distant things."

Many environmental phenomena commonly examined in archaeological studies, particularly distance measures, exhibit significant levels of spatial autocorrelation. To illustrate, I undertook a simulation study (Kvamme 1985b) that utilized simple random samples ($n = 100$) of 1 ha locations from a 100 km² region. The data were obtained from a working geographic information system (see Chapter 10), which facilitated the simulation. At each location, elevation and slope (percent grade), two commonly used variables in archaeological studies, were determined. An autocorrelation coefficient (Cliff and Ord's [1973] I statistic) was calculated for these variables for each of five simulation runs (where a new sample was selected for each run). In an associated significance test that yielded standard z-scores (a common statistic used to evaluate significance; see Thomas 1976), the z-scores ranged from 3.71 to 10.17 for slope (with an average of 6.15) and from 7.72 to 9.95 for elevation (with an average of 8.46). These scores indicate highly significant levels of spatial autocorrelation for these rather common variables, pointing to a lack of independence between the observations, since a z-score of 1.64 is significant at the 0.05 level and a z-score of 3.72 is significant at the 0.001 level (these tests were one-tailed). Thus, the real world presents difficulties even for simple random samples; researchers who somehow are able to obtain them and argue statistical correctness may be working with a false sense of security.

In regional archaeological analyses we often have no choice but to use some form of cluster sampling to obtain representative samples of sites from within a region. As Holmes (1970:381) states,

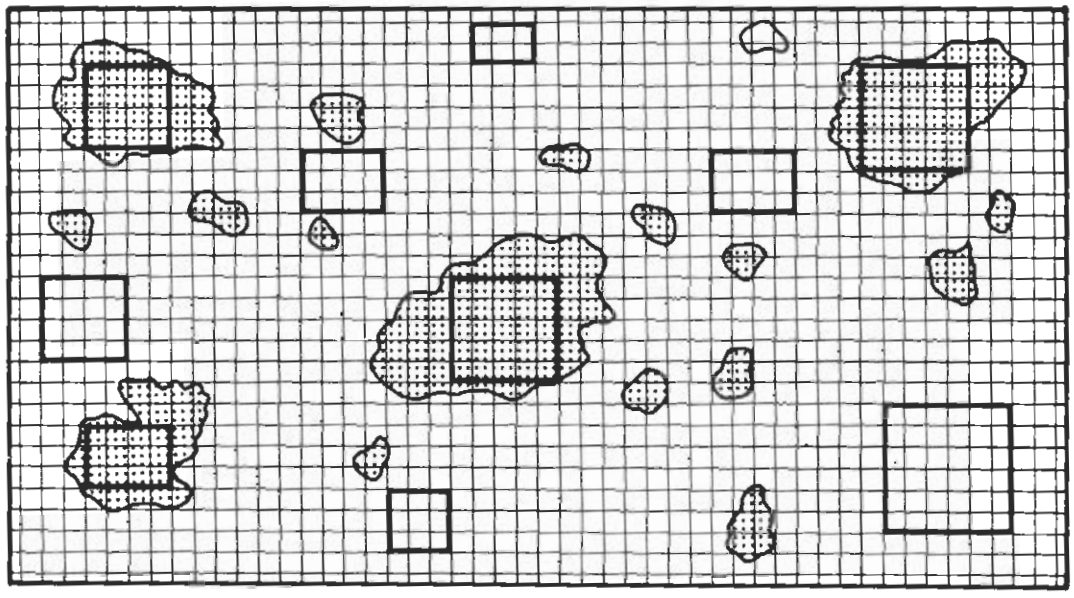
> this design sometimes must be employed because of *identification problems*. In some cases identification of individuals for sampling is absolutely impossible, while in others it is an extremely laborious, unrewarding task. These nonmathematical identification difficulties, arising from the nature of the data, will always persist in certain types of research, so that the need for area [cluster] samples will continue in these important fields [emphasis original].

Hence, it is important to examine the effects of cluster sampling on quantitative classification models *before* looking at the modeling approaches themselves, since many models of necessity are based on cluster-sampled data.

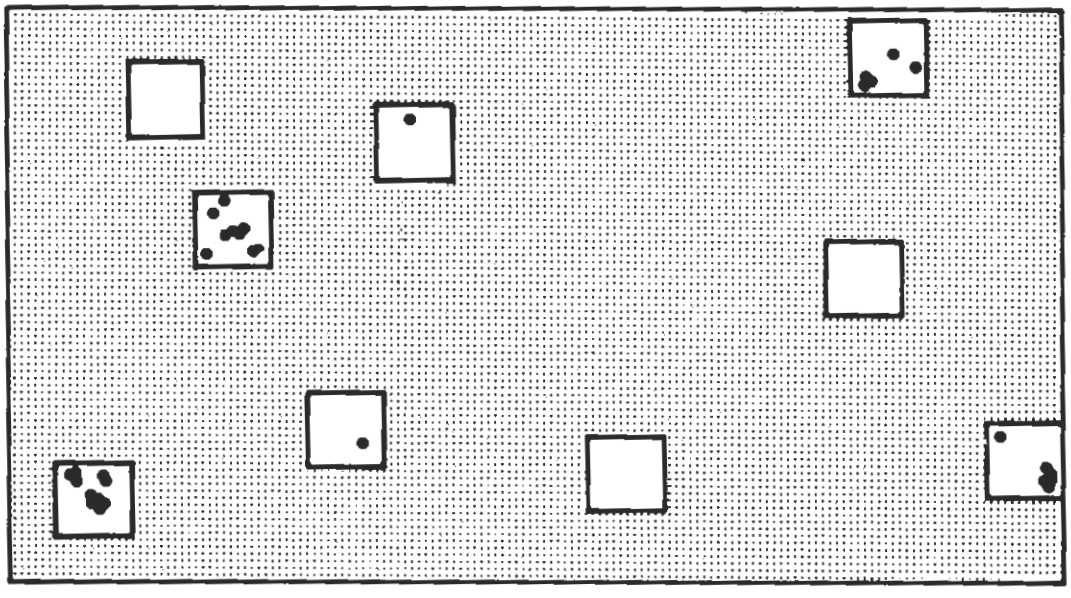
The typical sampling practice employed in image-analysis studies (Moik 1980:Fig. 8.7; Schowengerdt 1983:192, Fig. 3-30) is informative. This procedure is portrayed in Figure 8.6a, which illustrates forested areas (shaded) and nonforested areas (unshaded). The large blocks represent ground-truthed clusters of small cells or pixels of known group membership; in the remainder of the image, group membership is unknown. In each of the pixels, measurements of reflected radiation are recorded. A goal might be to develop a predictive model that classifies forest locations in the remainder of the image based upon reflectance characteristics of the known forest and nonforest pixels. This form of cluster sampling is somewhat more extreme than that typical in archaeological sampling, since the analysis elements (pixels) occur in tight, contiguous blocks (compare Figure 8.6a with Figure 8.6b, a typical archaeological sampling example). One might expect that in the Figure 8.6a example there would be a high degree of positive spatial autocorrelation because of the increased relative proximity between analysis elements. A second drawback of cluster-sampled data is that within-class variation tends to be underestimated (this follows from the reduced variability within clusters), making classes appear more different than they really are. The possible drawbacks of cluster sampling must be weighed against its benefits; in the remote-sensing case (as in archaeology), cluster sampling is much less difficult and costly than obtaining simple random samples of elements. As Schowengerdt (1983:192) states,

> In all random sampling procedures, it is desirable to select random *groups* of pixels rather than single pixels because of the practical difficulty in accurately locating single pixels on the ground [emphasis original].

In a classification perspective, the detrimental effects of well-designed cluster sampling seem to be small, as indicated by excellent classification results typically obtained by remote-sensing studies (Moik 1980; Schowengerdt 1983; Swain and Davis 1978). It is easy to see why this is true: a classification procedure only partitions the measurement space (Figure 8.5b). Differences between the measurement spaces defined by simple random samples and those defined by suitably constructed cluster samples are rather small when compared with differences between discriminable classes, particularly when the site-present and site-absent contrast discussed earlier is used. One simulation study (Campbell 1981) compared the performance of the dense cluster sampling practice illustrated in Figure 8.6a



A



B

Figure 8.6. Sampling practices. (A) In remotely sensed image studies analysis, elements or pixels (the small cells) of known group membership are typically selected in large contiguous blocks as a practical and cost-saving measure, here from forested (shaded) and nonforested (unshaded) settings. (B) A typical archaeological sampling practice consists of randomly placed quadrats with a variable number of sites (black dots) in each quadrat.

with simple random samples in classifying forested areas (vs nonforested areas) in several different Landsat scenes. The classification accuracy of the predictive (discriminant analysis) models obtained from the less autocorrelated simple random samples ranged from 15 percent better to 2 percent worse (an average of 6 percent better) than that of the corresponding models obtained from the more highly autocorrelated cluster samples. The lesson to be learned from this evidence is that we should not be too concerned with the detrimental effects resulting from the use of cluster samples, considering the benefits that are derived in return.

Notwithstanding these results, the question about the correctness of using statistical inferential procedures in these contexts still remains. Researchers faced with similar problems have developed robust model validation procedures in remote sensing and elsewhere (Schowengerdt 1983; Swain 1978). As described earlier, a site location model can be viewed simply as a classification or decision rule. For the moment let us forget how such a model might be developed. Rather, let us focus only on the idea that we have a decision rule, however it was derived, and that we can apply it to measurements obtained at locations (land parcels). Based on these measurements, the decision rule yields, at the very simplest, an assignment to one of two categories—for example, site present or site absent. If the decision rule has some predictive capacity in terms of the populations under study, then it should offer correct decisions more often than could be attributable to chance. This notion can be tested in practice by obtaining new random samples of known site-present and site-absent locations (both entirely independent of any samples that might have been used earlier in model development), by applying the decision rule to the measurements at these locations, and by determining how well these sites and nonsites are classified. If the percentage correctly classified is greater than that attributable to chance, then the decision rule has some predictive capacity, and it is here, in this testing phase, that methods of statistical inference are more appropriately applied. Relatively simple statistical testing procedures can be used to assess the significance of model classification results (see below).

The use of independent test samples makes this overall approach *robust* because performance is assessed on entirely new sets of data, which gives an excellent idea of true model performance in practice and obviates the need for reliance on the assumptions of multivariate statistical theory (e.g., multivariate normality and homogeneity of variance) in the model-development stage. Note that in this scheme it does not matter what procedures are used to develop a model or decision rule, *since all inferences about the usefulness of a model are drawn from the independent test samples*. Thus, any procedure can appropriately be used to formulate a decision rule, from simple subjective notions about site locations to complex multivariate data models. Regardless of the model-building procedure used, however, the statistical assessment of its worthiness is made through independent test samples. This is the approach taken in the remainder of this chapter. Such multivariate techniques as multiple discriminant analysis and logistic regression are used in subsequent sections for model development, but only as a means of obtaining a partitioning of the measurement space in the form of a decision rule. These

algorithms are based on very powerful mathematical differencing techniques that are able to provide excellent partitionings of the measurement space even when underlying assumptions are not fully met.

Example Analysis Based on Locational Characteristics

A site location study performed in the Glade Park region of western Colorado (Kvamme 1983c) can be used to illustrate model building based on locational characteristics. In this section these data are used to illustrate one approach to model development based on environmental and terrain characteristics observed at the known site and nonsite locations found by that study; a later section will use these data to illustrate model-testing procedures. For simplicity the model is developed for the locations of all open-air sites within Glade Park, although identical methods would apply for specific site-type model development (see "Modeling Individual Site Types," below). Only environmental factors are considered in this analysis. Variables measuring social factors would, if available, be treated in an identical manner, but in the present study, which dealt with hunter-gatherer archaeology, contemporaneity of the sites was impossible to establish from the survey data and such features as central-place settlements simply did not exist. The analysis was carried out by treating each hectare (100 by 100 m parcel) as the unit of investigation and then comparing land parcels that included sites with land parcels that did not.

The Glade Park study region, encompassing nearly 650 mi^2, lies on the western border of Colorado in the Bureau of Land Management's Grand Junction Resource Area. This arid region of mesa and canyon country is covered by piñon-juniper forests interspersed with grassy clearings and is archaeologically one of the richest areas of Colorado outside the southwestern portion of the state (Wormington and Lister 1956). The archaeological sites uniformly consist of small scatters of chipped stone artifacts, lithic debris, and occasional ground stone; ceramics are extremely rare.

Sampling

The purpose of the archaeological survey conducted in Glade Park was to obtain a random sample of site locations to be used in a modeling study of patterns of prehistoric site distribution. This was accomplished by surveying 38 quarter-sections randomly selected from a total of nearly 2600. These quarter-sections were gridded into 64 units of 1 ha each, which were the primary analysis units. Prehistoric sites were discovered in 157 of these 1 ha units out of a total of 2432 units examined. Of the 2275 land parcels that did not contain sites, a random sample of 157 was drawn to serve as the nonsite control group. It should be noted that, because nonsite locations were selected from a limited number of quarter-section clusters, environmental (and other) variation may have been underestimated, which can have a

deleterious effect on the performance of the resulting model. This practice of selecting nonsites from the same clusters as those in which sites are found also tends to make nonsite samples more similar to site samples than is really the case in nature, and this also can weaken a model. Although the Glade Park sample may not be optimal, it will be shown that very good results can be obtained through the nonsite sampling procedure used here.

An alternative nonsite selection approach that resolves these problems to some extent was used in a Colorado plains site location study (Kvamme 1984). This approach recognizes that in many regions archaeological sites are an extremely rare phenomenon, occurring by chance on the order of around 1 percent of the time (see the section below on "Base Rate Probabilities"). In other words, for every acre in a region that contains a site there might be 99 acres where no sites occur. The alternate approach draws a simple random (or other) sample of control locations from across the entire landscape of the region regardless of whether or not the locations have been field inspected for archaeological resources. The advantages of this approach are *(a)* that the resulting control group represents the true range of background environmental variation and *(b)* that levels of spatial autocorrelation are reduced (since selection is not by clusters). The disadvantage is that by chance a small percentage—in the above example around 1 percent—of the control locations actually falls on sites, which introduces error in group identification. This error is usually negligible and has little effect on analysis. A control group obtained in this manner may still be referred to as a "nonsite" group since under such conditions the vast majority of the group (99 percent in this example) really are nonsites. Obviously, in areas where the probability of finding a site is high this procedure should not be undertaken.

Environmental Variables

Fourteen environmental and terrain variables were measured at the center of each of the 157 site and 157 nonsite units. The variables were measures pertaining to landform, water, view, and shelter. The landform variables were slope measured as percent grade (Figure 8.2a) and local relief within 100, 250, 500, and 750 m (Figure 8.2b). The water variables consisted of horizontal and vertical distance to nearest stream and to nearest permanent river. The view variables were distance to nearest point of vantage and a measure of the angle of view (Figure 8.3a). The shelter variables consisted of aspect measured relative to north or south (using the 180° rescaling technique described above) and shelter volume measured within 100 m and 250 m (Figure 8.3b) but rescaled such that low (negative) values suggest relatively little shelter (hills) and high (positive) values suggest relatively high shelter (valleys).

Univariate Examination

The sample means, trimmed means (removing 15 percent of the largest and smallest values), medians, and standard deviations (Chapter 5, this volume) for the

site and nonsite samples are presented in Table 8.1. It is readily apparent that some major differences in environmental patterning exist between the site-present and site-absent (nonsite) groups. For example, sites tend to occur closer to water, on relatively level ground, and in regions of less local relief, and they tend to have better views. Sites also tend to occur under limited ranges of environmental variation (as indicated by somewhat smaller standard deviations).

Given such results, most researchers attempt to assess the statistical significance of the data patterning (e.g., Lafferty 1981; Larralde and Chandler 1981). The two-sample t-test (Thomas 1976:227) is a test for the difference between means, but use of this test requires such assumptions as normality and equal group variances. The Mann-Whitney test (Conover 1971:224) is a nonparametric alternative, and the Kolmogorov-Smirnov test can be used to assess distributional differences of any kind (Conover 1971:309). As noted in earlier sections, however, use of these tests in spatial contexts is problematic because of positive spatial autocorrelation, which violates the common assumption of independence. Since the Glade Park spatial data are derived from cluster sampling, we might expect the level of spatial autocorrelation to be rather high.

One way to resolve this difficulty is to treat such tests conservatively by using the 0.005 level instead of the 0.05 level, for example. When the t-test is used, the absolute value of t itself can serve as a relative index of difference or separability between classes. Currently there are no readily available significance tests for assessing class differences in spatially autocorrelated contexts (however, see Cliff and Ord 1975).

A modified t-test valid for unequal group variances (Steel and Torrie 1980:206) was applied to the Table 8.1 data using a robust procedure that trims the largest and smallest values in each group (Dixon et al. 1983:101), since the t-test is overly sensitive to extreme scores. The t-statistics and associated two-tailed probabilities are given in Table 8.1 and are presented only as a relative index of separability between the site and nonsite groups. The t index suggests that certain variables, such as slope, aspect, and view, are more separated than other variables. Statistically, the results of the t-test should be viewed conservatively because of violations of the independence assumption that result from spatial autocorrelation. Additionally, even if the data could be assumed to be independent, the resulting statistics would still be inflated because simultaneous inference methods (Miller 1966) were not employed. Besides correlation between *cases* resulting from spatial autocorrelation, the 14 *variables* are also positively correlated (e.g., horizontal and vertical distance to water are highly related). Thus, the 14 individual significance tests are not independent assessments; moreover, with 14 tests some are likely to appear significant by chance alone.

Multivariate Assessment

Before attempting to model the site and nonsite differences that appear to exist in the Glade Park data (Table 8.1), it might be instructive to assess group

TABLE 8.1.

Descriptive statistics for Glade Park sites and nonsites

Variable		*Mean*	*Trimmed Mean*	*Median*	*s.d.*	*t-test t/p*
X_1	slope (% grade)					
	sites	12	10	10	11	-5.28/0.0000
	nonsites	24	18	13	26	
X_2	relief within 100 m (m)					
	sites	30	27	24	20	-3.00/0.0030
	nonsites	39	34	30	30	
X_3	relief within 250 m (m)					
	sites	70	65	61	45	-2.65/0.0085
	nonsites	85	79	73	52	
X_4	relief within 500 m (m)					
	sites	133	130	134	63	-3.11/0.0021
	nonsites	160	146	146	90	
X_5	relief within 750 m (m)					
	sites	180	180	183	78	-3.33/0.0010
	nonsites	217	200	195	112	
X_6	horizontal distance to permanent water (m)					
	sites	2273	1943	1950	1960	-0.75/0.4566
	nonsites	2426	2254	2300	1689	
X_7	vertical distance to permanent water (m)					
	sites	129	98	91	138	-2.16/0.0319
	nonsites	168	125	125	175	
X_8	horizontal distance to nearest water (m)					
	sites	164	139	100	133	-1.94/0.0535
	nonsites	194	183	200	136	
X_9	vertical distance to nearest water (m)					
	sites	33	26	24	37	-2.22/0.0270
	nonsites	44	33	24	48	
X_{10}	vantage (m)					
	sites	159	68	50	277	1.14/0.2556
	nonsites	130	86	100	182	
X_{11}	view angle (0–360°)					
	sites	219	222	220	73	5.51/0.0000
	nonsites	174	178	180	69	
X_{12}	aspect (0–180°)					
	sites	64	57	50	48	-4.15/0.0000
	nonsites	89	87	90	55	
X_{13}	shelter within 100 m					
	sites	-14	-9	-6	43	-2.94/0.0035
	nonsites	3	-1	-3	57	
X_{14}	shelter within 250 m					
	sites	-140	-93	-59	546	-1.63/0.1037
	nonsites	-28	-78	-73	654	

differences by considering all available information (the 14 variables) simultaneously. Hotelling's T^2, a multivariate extension of the t-test, and one-way multivariate analysis of variance (MANOVA) are traditional parametric procedures for performing such a task (Morrison 1976). Recently, a nonparametric alternative has been presented for a similar problem context in archaeology. Multi-Response Permutation Procedures (MRPP) originally were introduced to archaeology for assessing artifact class locational differences in *real* space based on positional coordinates (Berry et al. 1980, 1983, 1984). MRPP can be used in the present situation to assess multivariate site and nonsite class locational differences in *measurement* space (Figure 8.5b). Since MRPP are based on a randomization procedure, they are extremely robust. If substantial class differences are found, this result would suggest that the site and nonsite locations occupy different regions of the measurement space. Site location modeling procedures then might have a reasonable chance of partitioning the measurement space into appropriate decision regions, providing a successful classification model.

The Glade Park site and nonsite locational data were subjected to an MRPP analysis. The simultaneous comparison of all 14 site and nonsite environmental characteristics indicates an extreme difference between the two classes that was significant at $p = 0.00000000032$. This suggests that the Glade Park locations with sites tend to be markedly different from locations without sites in terms of environmental characteristics (see "Interpretation and Explanation of Data Patterns" for a discussion of how such data patterns can be interpreted).

Site Location Models

The technique chosen for site location model development at Glade Park is multiple logistic regression. This classification algorithm is particularly robust because, unlike many other classification techniques, it does not assume a particular underlying distributional form (Press and Wilson 1979) but achieves a partitioning of the measurement space (Figure 8.5b) based on the empirical distribution of the particular data set used (see Chapter 5 and the discussion below for more details about logistic regression). The following logistic regression equation was obtained through the BMDP program LR (Dixon et al. 1983):

$$L = 0.713 - 0.0390X_1 - 0.00454X_2 + 0.00602X_3 + 0.00526X_4 - 0.00540X_5 - 0.0000828X_6 - 0.00186X_7 - 0.000628X_8 - 0.0126X_9 - 0.00808X_{10} + 0.00748X_{11} - 0.00519X_{12} - 0.0178X_{13} + 0.000746X_{14}$$

where the variables referred to by the X_i may be found in Table 8.1. The value of L theoretically can range between positive and negative infinity; positive values denote locations in the site portion of the measurement space, negative values indicate locations in the nonsite portion, and $L = 0$ represents locations that fall exactly on the decision boundary (Figure 8.5b). Thus, L represents a decision rule that can be used to assign locations to site or nonsite classes on the basis of their measurements. Additionally, large positive or negative values indicate locations having characteristics that, overall, are more like the site or nonsite classes,

respectively, than locations with small positive or negative values. L therefore represents a single scale or axis representing an underlying environmental continuum with "site-favorable" conditions on the positive extreme and "site-unfavorable" (nonsite) conditions on the negative extreme.

In practice, the use of L is unwieldy because its values are unconstrained. A simple transformation yields a value that ranges from 0 (large negative values of L) to 1 (large positive values of L), with 0.5 indicating locations on the decision boundary ($L = 0$).

$$p_i = \frac{e^{(L_i)}}{1 + e^{(L_i)}} = \frac{1}{1 + e^{(-L_i)}}$$

where L_i is the logistic regression score measured at the ith location and p_i is the transformed value. Note that if *(a)* the data are obtained through simple random sampling (generally impractical in archaeology, as noted above), and *(b)* the data represent independent observations (generally impossible owing to the spatially autocorrelated nature of archaeological data), then the p_i values can be interpreted as estimates of a location's probability of membership in the site-present class conditional on the measurements (X_i) made at the location. Since these two conditions are not met in the Glade Park analysis, the p_i values can best be interpreted as standardized relative indications of location within the site-present or site-absent portions of the measurement space.

To illustrate use of these formulas, suppose that a location is found to exhibit measurements on 14 predictor variables identical to those presented for the site group mean values in Table 8.1. When the site mean values and the first equation are used, $L = 0.6331$; the second equation gives $p = 0.6532$. Thus, a location with those environmental characteristics would be assigned to the site-present group since $p \geq 0.5$. When this procedure is applied to the measurements of all 314 site and nonsite locations, the initial classification results are as shown in Table 8.2a. The percent correct statistics in this table are undoubtedly inflated because the same data were used both to build the model and to yield these performance indications (a model like logistic regression tends to maximize fit to the particular data at hand). In a later section, independent tests are applied in an attempt to assess the "true" performance of a Glade Park model. The gain statistic for this model is 0.60.

One problem in applying a model such as the one presented in the above equation is that measuring many variables and performing many calculations requires much work, even for a computer. A common data reduction technique is principal components analysis (Morrison 1976) by means of which the variation in a large number of variables is typically summarized by a smaller number of dimensions (principal components), which are linear combinations of the original variables. This technique is also used to eliminate redundancies resulting from intercorrelations (collinearity) among variables. The reduced number of components can be used as predictor variables in classification analyses (Schowengerdt 1983:160).

TABLE 8.2.

Classification performance of initial Glade Park models

A. 14-Variable Model

		Predicted Group	
		Site	*Nonsite*
Actual Group		$p \geq 0.5$	$p < 0.5$
Site	Number	107	50
	Percent	(68.2)	(31.8)
Nonsite	Number	43	114
	Percent	(27.4)	(72.6)
		Gain = 1 - (27.4/68.2) = 0.60	

B. 9-Variable Model

		Predicted Group	
		Site	*Nonsite*
Actual Group		$p \geq 0.5$	$p < 0.5$
Site	Number	110	47
	Percent	(70.1)	(29.9)
Nonsite	Number	53	104
	Percent	(33.8)	(66.2)
		Gain = 1 - (33.8/70.1) = 0.52	

Principal components analysis has not been used extensively in site location model development, principally because it is very difficult to interpret the meaning of the components obtained. Moreover, in order to obtain component scores for each case (location) to which the model might be applied, the technique requires measurements of the original variables anyway, and thus there is little savings in time and effort.

Various stepwise procedures present an alternative (see Chapter 5). These techniques attempt to select a "best" subset of variables for a model. *Best* in this case means that the addition of other variables will not substantially improve the model because they contain only redundant information (owing to intercorrelations). In forward stepping, the first step selects the single variable that offers the maximum discrimination between groups (indicated by some statistic) and enters this variable into the model. The second step selects the variable from the remaining pool of variables that offers the greatest increase in discrimination between classes by considering the relationship of this second variable with the variable already in the model and with variables not yet in the model. On each succeeding step additional variables are selected and entered into the model until the remaining variables (those not yet entered) are determined to contain only redundant information (again owing to correlation between them and the variables already in

the model). The result is a subset of available variables that yields a model whose performance may be similar to that of a model in which all variables were used but that requires less information. One drawback of stepwise procedures is that the final subset of variables obtained in a particular application can vary depending on the particular stepwise procedure and the selection criteria used and can also vary from sample to sample. It is usually the case, however, that a certain core of best discriminating variables is selected.

The 14 variables of the Glade Park data were subjected to stepwise procedures using the BMDP stepwise logistic regression program (LR; Dixon et al. 1983). Variables were entered at each step on the basis of largest chi-square value (suggesting best discrimination). A subset of nine variables was ultimately selected by graphing at each step changes in several statistics, including *(a)* the improvement chi-square, *(b)* the model log likelihood, and *(c)* the goodness-of-fit statistie R_p^2 (described above), all of which are monotonically related functions. After the ninth variable was entered, changes in all of these statistics leveled off, suggesting that no substantial model improvement would occur if the remaining five variables were included. The resulting nine-variable model is

$$L = -0.158 - 0.0401X_1 + 0.518X_2 - 0.00224X_7 - 0.0133X_9 - 0.000602X_{10} + 0.00738X_{11} - 0.00582X_{12} - 0.0183X_{13} + 0.000804X_{14}$$

and for this model gain equals 0.52. The classification perfotmance of this nine-vatiable model, when it was applied to the same data used to create it, is shown in Table 8.2b. (Independent tests of Glade Park models are given below.) A comparison of the nine-variable model shown here and the 14-variable model described above shows that not only are the models very similar (in terms of the coefficients) but so are their suggested performances, as indicated by the statistics in Table 8.2.

APPLICATION COMPARISON OF QUANTITATIVE LOCATIONAL MODELS

In this section several forms of quantitative data models of characteristics of locations that have been used in archaeological research are presented and compared. Each type of model achieves a partitioning of the measurement space (Figure 8.5b) in a different way. The goal of this section is to demonstrate the broad similarity of these diverse modeling techniques and of their results. For comparison, each modeling technique is mapped across the same study region using GIS computer mapping techniques (see Chapter 10); rhe patterns of the mappings are often strikingly parallel. The results of this section support the conclusion alluded to earlier and arrived at by Hixon et al. (1980): the particular classification algorithm used is *less* important than the representativeness of the samples used in predictive model development.

All models presented in this section were developed using the same data from a Colorado plains study region (Kvamme 1984, 1986). This study region of nearly 575

km² was gridded into approximately 230,000 cells (land parcels), each measuring 50 m on a side; these cells were the elementary units of analysis. The archaeological data used for the models consisted of 269 locations (cells) containing open-air lithic scatter sites. A background control group of 1154 locations without archaeological remains (nonsites) was also used. These large sample sizes should help to illustrate the relative performance of each modeling procedure. Eight environmental variables calculated at these locations by a computer through GIS techniques formed the data base, and in all models all eight variables are used for consistency in comparison. The variables are aspect (X_1), slope (X_2), local relief within 100 m (X_3), local relief within 300 m (X_4), a canyon rim index value (the "shelter" volume measure described above; X_5), distance to nearest point of vantage (mesa edge, canyon rim, or hill or ridge top; X_6), distance to the closest drainage (X_7), and distance to the closest second-order (or greater) drainage (using Strahler order ranking; X_8). The reader is referred to the section "Variables Used In Locational Research" for a description of how these variables are measured.

Robust Classification Models

Robust classification models can be grouped according to two types—*parametric* and *nonparametric*. Parametric techniques assume a particular type of statistical distribution (usually multivariate normality) and then estimate parameters of that distribution (e.g., means, variances, and covariances). Nonparametric classification procedures make no assumptions about distributional form and are sometimes considered particularly robust because they work under a wide range of distributional types (if the groups to be classified are reasonably distinct). It should be noted that under the same conditions (varied distributional types) parametric methods usually provide good results, even when the assumed (multivariate normal) distribution does not occur (Schowengerdt 1983:176).

Two parametric techniques, a linear discriminant function (commonly called discriminant analysis) and Bayes's maximum likelihood, are compared below with logistic regression, a nonparametric technique. The section ends with a brief discussion of a quadratic classification technique.

Discriminant Analysis

Because many archaeologists are familiar with discriminant analysis and because the software needed to use this technique is readily available in common statistical packages, it has been the dominant technique used for site location model development (Kvamme 1980; Larralde and Chandler 1981; Peebles 1981; Schroedl 1984).

The overall strategy of discriminant analysis in the two-group situation entails summarizing class differences by a linear combination of the original and multiple variables, where each observation is assigned a score on the single resulting

dimension or *discriminant axis*. The discriminant function has the characteristic of maximizing the separation between groups along the axis, assuming multivariate normality and equal group variation. A maximum likelihood technique is then used on the discriminant axis to evaluate probabilities of group membership (see Chapter 5).

When the example data of 269 site and 1154 nonsite locations are used, the following discriminant function is obtained through the BMDP discriminant analysis program 7M (Dixon et al. 1983).

$$D_i = -5.7058 - 0.0047X_{1i} - 0.08X_{2i} + 0.05X_{3i} - 0.0164X_{4i} + 0.00131X_{5i} - 0.0002X_{6i} - 0.001X_{7i} - 0.0008X_{8i}$$

where D_i is the discriminant score for the ith case (location) and X_1 through X_8 are the variables defined above. Like L, D can range between positive and negative infinity. A simple transformation yields a value p, which ranges from 0 to 1, allowing interpretation (when the assumptions of this model are fully met) as the probability of a location's membership in the site group, conditional solely on the measurements. This transformed value is calculated as follows:

$$p_i = \frac{e^{-0.5\,(D_i - D_s)^2}}{e^{-0.5\,(D_i - D_s)^2} + e^{-0.5\,(D_i - D_{ns})^2}}$$

where D_s is the estimated mean (centroid) of discriminant scores for the site group and D_{ns} is the mean discriminant score for the nonsite group.

To illustrate use of these formulas, the environmental characteristics of site 5LA5364, one of the 269 sample sites, are shown in Table 8.3. When these data are used, the first equation yields a discriminant score of $D = 2.1909$. If this value and the site and nonsite sample centroids on the discriminant axis ($D_s = 0.8304$, $D_{ns} = -0.1936$) are inserted in the second equation, then $p = 0.8719$. Since p is greater than 0.5, which is the traditional decision rule, this location would be correctly classified as a site. Of course, this model can be used for prediction when it is applied to locations of unknown group membership.

Replicating this procedure for the 1423 site and nonsite sample locations yields the initial model accuracy indications (percent correct statistics) shown in Table 8.4. The gain statistic can be estimated from these data. The percentage of the total area covered by the model (at the $p = 0.5$ cutoff) can be estimated by using the percentage of nonsite locations classified by the model to the site group, 32.1 percent in this case. We can use this figure as an estimate because in the region under study (as with most regions) the area occupied by site locations constitutes only a tiny percentage of the study area—in the present case about 1 percent of all the locations (50 m cells) in the region, which means that the nonsite locations represent about 99 percent of the total area (see Kvamme 1984). Thus, if the model classifies 32 percent of the nonsites to the site group, we can infer that approximately 32 percent of the total area of the region would be classified in the site group

TABLE 8.3.

Values for eight environmental variables: 5LA5364, all 269 sites, and 1154 nonsite locations in the Colorado plains study region

		5LA5364	*All Sites*		*All Nonsites*	
Environmental Variables		X_i	$\bar{x}$	*s.d.*	$\bar{x}$	*s.d.*
X_1	aspect (°)	31.0	79.8550	52.7906	97.196	52.4157
X_2	slope (% grade)	5.0	4.3090	4.1598	4.452	8.7874
X_3	relief within 100 m (m)	24.4	13.5750	10.4165	11.263	17.4562
X_4	relief within 300 m (m)	30.5	27.6383	18.3213	24.705	27.2667
X_5	canyon rim index ($m^5/1000$)	6104.0	5908.1320	460.7342	5731.842	392.0943
X_6	vantage distance (m)	72.0	472.5350	637.6297	916.471	841.7876
X_7	distance to closest drainage (m)	144.0	147.5910	165.4219	365.180	366.7622
X_8	distance to second-order drainage (m)	144.0	392.7550	465.3968	794.431	602.3548

by the model if it were mapped. The percentage of all sites within the model's area is estimated by the percentage of sites correctly classified by the model to the site group—76.2 percent for the current model (Table 8.4). Based on these calculations, the discriminant model yields a gain of 1 - (32.1/76.2) or 0.579. It should be noted that both the percent correct and the gain statistics presented in Table 8.4 are inflated owing to a variety of factors, the most notable of which is that the same data were used both to build the model and to evaluate its performance (procedures given below help to correct inflated statistics through independent tests).

We can illustrate the application of this model when mapped over a region by using the Colorado plains 5.5 by 8.5 km tesr study region discussed in the section entitled "Approaches Based on Trend in Location Only." This region, which can be characterized as a level plain dissected by a number of deeply entrenched canyons, represents only a small portion of the larger study area from which the sample data were derived. Approximately half of the 95 sites in this test study region are contained in the larger sample of 269 sites used for development of this model. Computer measurement and mapping techniques in the form of a geographic information system (see Chapter 10) were used to estimate values for the eight variables in each 50 by 50 m cell and to map the results of the model over the approximately 19,000 cells of this test region (Figure 8.7a).

In this figure, and in those that illustrate mapping of the subsequently discussed models, estimated probability values are portrayed in five levels (in steps of 0.2 and in levels of increasing darkness). Thus, the traditional $p = 0.5$ decision rule lies midway within the second level of shading. The actual site locations in this study region were shown in Figure 8.4a and may be compared with this predictive

TABLE 8.4.

Comparison of classification performance rates of several site location modeling procedures (row percents are given in parentheses)

Discriminant Analysis Predictions			*Maximum Likelihood Predictions*			*Logistic Regression Predictions*		
	Site	*Nonsite*		*Site*	*Nonsite*		*Site*	*Nonsite*
Actual	$p \geq 0.5$	$p < 0.5$	Actual	$p \geq 0.5$	$p < 0.5$	Actual	$p \geq 0.5$	$p < 0.5$
Site	205 (76.2)	64 (23.8)	Site	202 (75.1)	67 (24.9)	Site	222 (82.5)	47 (17.5)
Nonsite	370 (32.1)	784 (67.9)	Nonsite	382 (33.1)	772 (66.9)	Nonsite	399 (34.6)	755 (65.4)
gain = 0.579			gain = 0.559			gain = 0.581		

Euclidean Distance (d2) Predictions			*City Block Distance (d1) Predictions*			*Level Slice (±1.75 s.d.) Predictions*		
	Site	*Nonsite*		*Site*	*Nonsite*		*Site*	*Nonsite*
Actual	$p \geq 0.5$	$p < 0.5$	Actual	$p \geq 0.5$	$p < 0.5$	Actual	$p \geq 0.5$	$p < 0.5$
Site	217 (80.7)	52 (19.3)	Site	218 (81.0)	51 (19.0)	Site	186 (69.1)	83 (30.9)
Nonsite	482 (41.8)	672 (58.2)	Nonsite	498 (43.2)	656 (56.8)	Nonsite	500 (43.3)	654 (56.7)
gain = 0.482			gain = 0.467			gain = 0.373		

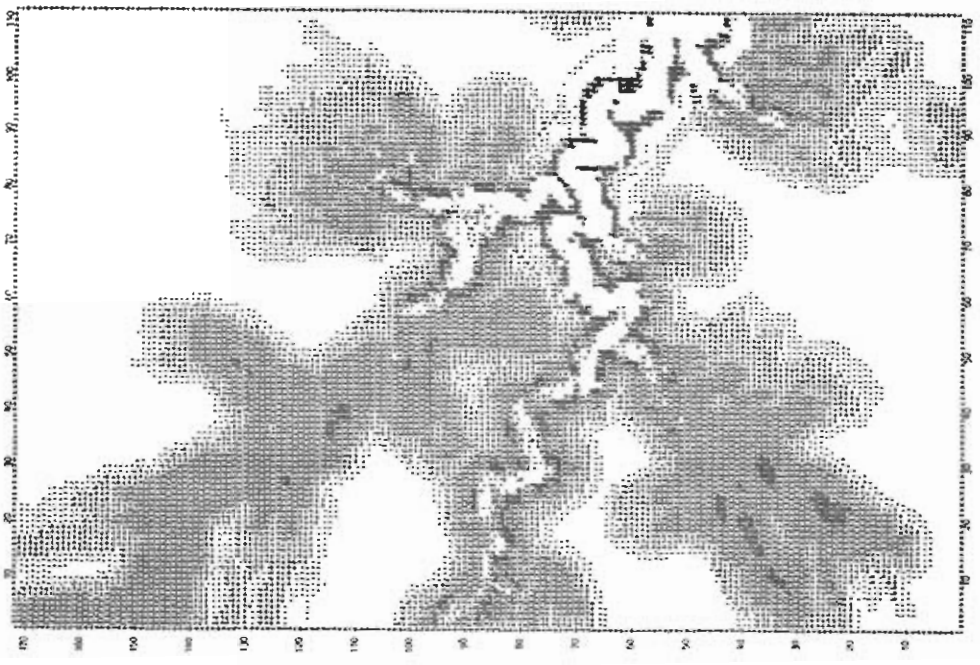
map. Since discriminant analysis assumes equal class variation, achieved by pooling the sample covariance matrices, a greater proportion of the environment tends to be classified with higher p-values (as indicated by the extent of the shaded regions) with this technique than with other methods described below that do not make this assumption (compare Figures 8.7b and c).

Maximum Likelihood Classifier

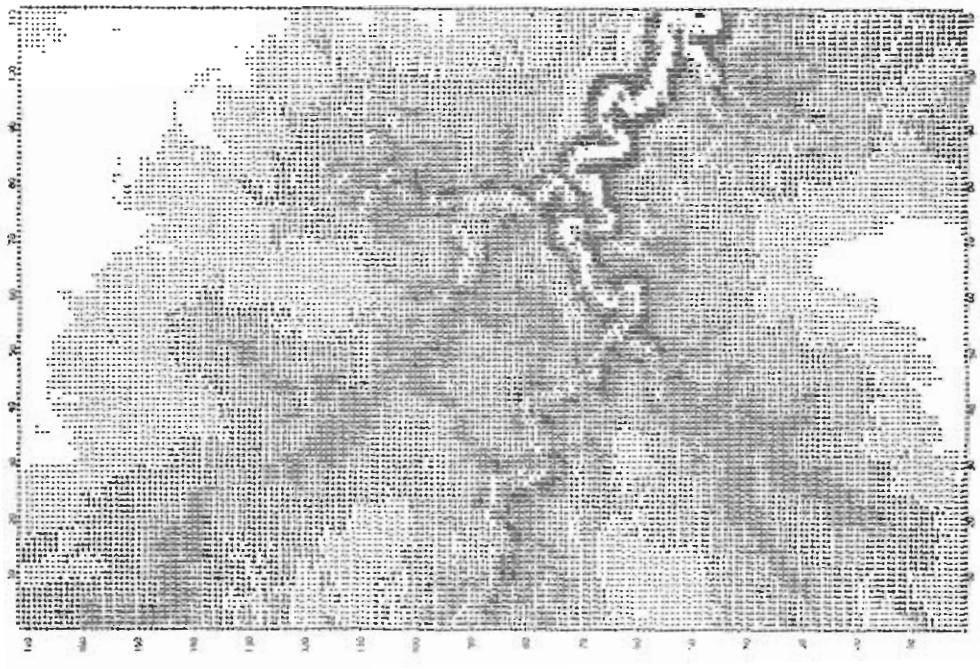
The maximum likelihood classifier is the most commonly used classification procedure in many disciplines, particularly in remote-sensing modeling applications (Moik 1980; Schowengerdt 1983), although it has been used less frequently in archaeological site location studies (Morain et al. 1981).

The probability that an observation belongs to the kth class, according to multivariate normal theory, is described by the following function:

$$\Phi_k(\mathbf{X}_i) = \frac{1}{(2\pi)^{p/2}|\Sigma_k|^{1/2}} e^{[-1/2(\mathbf{X}_i - \mu_k)' \Sigma_k^{-1} (\mathbf{X}_i - \mu_k)]}$$



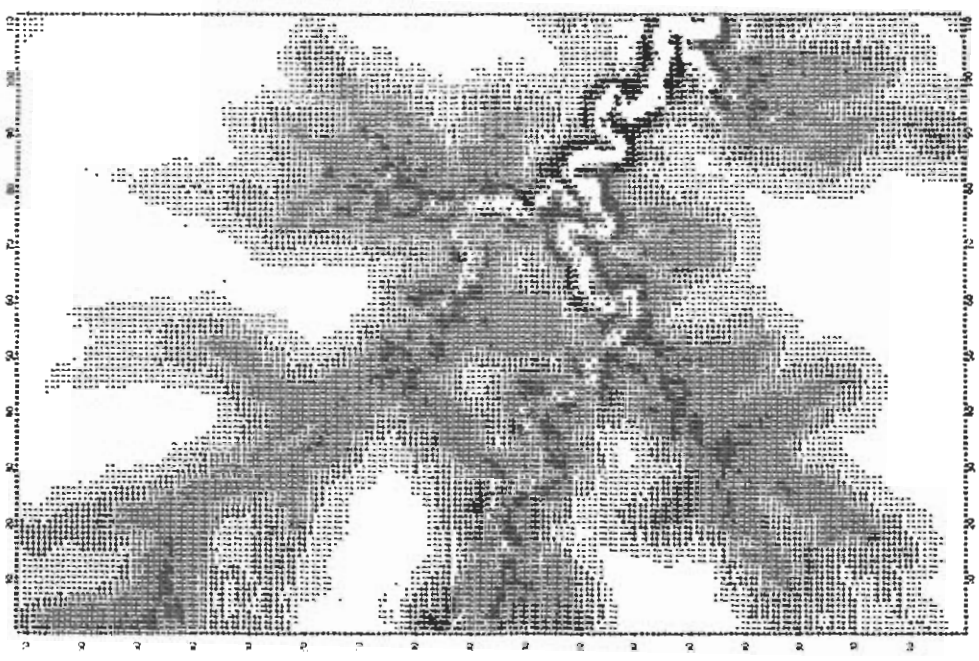
B



A

Figure 8.7. Robust classification models based on characteristics of locations (eight environmental variables): (A) discriminant analysis, (B) maximum likelihood, (C) logistic regression.

(continued)



C

Figure 8.7. Continued.

where $\mathbf{X}_i$ refers to vector of measurements of the multiple variables at location i, μ_k contains the vector of multiple mean values associated with class k, and Σ_k is the corresponding dispersion matrix containing j rows and columns of variances and covariances for class k (Swain 1978:156; also see Green 1978 for a discussion of matrix algebra techniques). In practice, the means and dispersion matrices are unknown; they are estimated by sample means, variances, and covariances. An observation is assigned to the class for which it has the greatest probability value.

Although discussion of matrix algebra is beyond the scope of this presentation, a simplified description of the technique follows. For a single variable we can imagine a normal probability curve with its maximum height or *density* at the mean value and with a width that is indicative of the variation in the distribution. For any value of a variable we can determine the density (height) of the distribution. Similarly, in a multivariate context multiple measurements can be assessed by the above formula relative to the multiple mean values for a class, considering at the same time the nature of the dispersion within that class, and this yields a multivariate normal density value. A density can be determined for each class under consideration. To illustrate with hypothetical values, if the multivariate density for Class A is determined to be 0.3 for the multiple environmental measurements made at some location and the density for Class B is determined to be 0.2 (in a two-class problem), then the measurements have a higher probability of belonging to Class A than to Class B; in fact, the probability of membership in Class A can be estimated as $p = 0.3/(0.2 + 0.3) = 0.6$. The mathematics of this procedure perform optimally when multivariate normality and independent observations can be assumed (i.e., classification error is minimized), even though this technique does not require equal covariance matrices (see Chapter 5). The Statistical Analysis System PROC DISCRIM performs multivariate classification through the maximum likelihood method (SAS Institute 1982).

To illustrate application of this technique, data from site 5LA5364 are again used (Table 8.3). Entering these data into the above equation, together with estimated means, variances, and covariances for the site and nonsite groups (some of which are included in Table 8.3), yields a density for the site group of 4.8539×10^{-21} and a density for the nonsite group of 3.4153×10^{-22}. Thus, the measurements at 5LA5364 indicate that this location has a higher probability of membership in the site-present than the site-absent group, and it would be appropriate based solely on these measurements to assign the 5LA5364 location to the site class. The conditional probabilities become

$$p\,(\text{site} \mid \mathbf{X}_i) = \frac{4.8539 \times 10^{-21}}{4.8539 \times 10^{-21} + 3.4153 \times 10^{-22}} = 0.9343$$

$$p\,(\text{nonsite} \mid \mathbf{X}_i) = 1 - p\,(\text{site} \mid \mathbf{X}_i) = 1 - 0.9343 = 0.0657$$

Applying identical calculations to the 1423 sample locations yields the initial model accuracy rates and gain statistic shown in Table 8.4. The results of the maximum likelihood technique applied to each of the 19,000 locations in the test study region are mapped in Figure 8.7b. Although the classification accuracy is about the same as that provided by the discriminant analysis, note that the maximum likelihood procedure maps a relatively smaller portion of the region as site likely because it takes into account the lesser environmental variation usually exhibited by a site-present class while the discriminant analysis model used above does not.

Logistic Regression

Multiple logistic regression has recently come into use as a classification technique (e.g., Maynard and Strahler 1981; Pindyck and Rubinfeld 1976:237–263; Schmidt and Strauss 1975), and it has been applied in several studies of archaeological site location (Custer et al. 1983, 1986; Holmer 1982; Kvamme 1983b, 1983c, 1986; Parker 1985; Scholtz 1981). This *nonparametric* technique makes no assumptions about distributional form (Wrigley 1976) and has been shown to offer improved classificatory performance over discriminant analysis when the data are not multivariate normal (Maynard 1981; Press and Wilson 1979). Maynard and Strahler (1981) argue that logistic regression is the optimal statistical classifier for remotely sensed data, and because no distributional assumptions are made, this technique is appropriate for nominal-, ordinal-, or interval-scaled data or for various combinations of these levels of measurement. (Logistic regression has been applied in several examples in earlier sections of this chapter.)

Logistic regression can be better understood if we consider the results of applying a multiple linear regression model to a dichotomous dependent variable, such as site presence and site absence. Such a model has a number of serious problems in this context (Wrigley 1976:7–9; Chapter 5, this volume), not the least of which is that predictions can range in value between plus and minus infinity, making it difficult to interpret these predictions as probabilities. Logistic regression is able to overcome these difficulties and yield a result that is constrained between 0 and 1. This result can be interpreted as a predicted probability of class membership (when assumptions of independence and random sampling are met) through use of the logistic transformation.

$$p_i = \frac{e^{(L_i)}}{1 + e^{(L_i)}} = \frac{1}{1 + e^{(-L_i)}}$$

where the logistically derived discriminant function at the ith location is

$$L_i = \alpha + \beta_1 X_{1i} + \beta_2 X_{2i} + \ldots + \beta_j X_{ji}$$

and α and the βj are the estimated intercept and regression weights.

A logistic regression analysis was performed using the example data of 1423 site-present and site-absent locations and the BMDP program LR (Dixon et al. 1983), and the result was the following function:

$$L_i = -6.8837 - 0.0043X_{1i} - 0.114X_{2i} + 0.0277X_{3i} - 0.0136X_{4i} + 0.00164X_{5i} - 0.000626X_{6i} - 0.0043X_{7i} - 0.000777X_{8i}$$

When applied to the measurements from 5LA5364 (Table 8.3), these equations yield $L = 1.9085$ and $p = 0.8708$. Based on its environmental characteristics, 5LA5364 would be correctly classified to the site-present group.

Model accuracy for the logistic regression application, as measured by the gain statistic in Table 8.4, is slightly higher than it was for the previous, parametric techniques. Figure 8.7c shows the results of mapping the logistic regression model over the test study area. Since logistic regression makes no assumptions about distributional form, it is usually regarded as a very robust procedure. This would appear to be an advantage for archaeological locational modeling because site location data are decidedly nonnormal, but in fact, the application of this technique to the sample data produced results that are very similar to the results obtained by the previous classifiers, both in performance (Table 8.4) and in mapped results (Figure 8.7).

Quadratic Classification Procedure

The quadratic classification method is a general technique that can be applied to such statistical models as discriminant analysis and logistic regression when group variances and covariances have been found to be markedly unequal. This procedure has been shown to offer improved classificatory performance in these situations (Anderson 1975; Eisenbeis and Avery 1972:44; Michaelis 1973; Smith 1947). In archaeological predictive modeling, Kohler and Parker (1986) have applied quadratic discriminant analysis to simulated data, and I have applied quadratic logistic regression as a test case in actual model development (Kvamme 1983c). The quadratic procedure incorporates all quadratic terms (i.e., squared terms for each variable and all possible interaction pairs) into a model, along with the predictor variables being used. This causes the decision boundary to wrap or curve around the group with less variation (a hypothetical quadratic decision boundary is shown in Figure 8.5b), which can provide an increase in model accuracy.

Any benefits obtained are not without cost, however. The discriminant analysis and logistic regression models presented in the previous sections require estimates of $j + 1$ parameters (where j is the number of predictor variables) to yield a linear decision boundary in the measurement space (Figure 8.5b). The nonlinear decision boundary of the quadratic model (Figure 8.5b) requires estimates of $(j + 1) + j(j + 1)/2$ parameters (thus a nine-variable model would require estimates of $[9 + 1] + 9[9 + 1]/2 = 55$ parameters). This increase in the number of parameters may require a corresponding increase in sample size in order for estimation to be reliable. Another problem is that, like the polynomial regression technique discussed in the

section on "Approaches Based on Trend in Location Only," the quadratic procedure does not produce a model that can be readily interpreted. Finally, in an application of the technique to archaeological site and nonsite data, I found it to be overly sensitive to outliers, which offsets most of the advantages gained through the inclusion of the extra terms (Kvamme 1983c).

Some Simple Classification Models

The models of the previous section constitute one set of approaches for partitioning the measurement space (Figure 8.5b) to achieve classification. When appropriate theoretical assumptions (such as multivariate normality) are met for each of these models, classification error in the partitioning that is obtained is minimized. As noted in earlier sections, however, many of these assumptions are difficult to meet when one is dealing with geographically distributed phenomena.

A number of simple mathematical rules have been developed to achieve a partitioning of the measurement space in pattern-recognition and image-analysis studies (Duda and Hart 1973; Moik 1980; Schowengerdt 1983). These procedures can be classed as nonparametric because no assumptions are made about probability distributions, and in some cases they perform with accuracy rates comparable to those of the models discussed in the previous section. An important advantage of these procedures is that they are easier to calculate (many can be done by hand) than the computationally burdensome procedures described above. Although a wide range of possible examples exist, only two are discussed here: the minimum distance classifier and the level slice classifier.

Distance Measures

The ***minimum distance*** algorithm simply classifies a location to the class that it is "closest" to in the measurement space (Figure 8.5b). In other words, a location (with characteristics summarized by measured variables) is assigned to one of the classes if its distance from the center of that class is less than its distance from the center of the other class(es) (Schowengerdt 1983:49–53). The center of each class is represented by the point in the measurement space having the class mean value for each variable under examination. The distance from the kth class is given by

$$d2_k = [(x_{1i} - \mu_{1k})^2 + (x_{2i} - \mu_{2k})^2 + \ldots + (x_{ji} - \mu_{jk})^2]^{1/2}$$

which is simply the Euclidean distance between the values of the j variables ($x_1, x_2, \ldots, x_j$) measured at the ith location, and the mean values for each variable ($\mu_1, \mu_2, \ldots, \mu_j$) for the kth class.

To illustrate application of this algorithm, the measurements for 5LA5364 and the estimated means for the sample site (s) and nonsite (ns) classes (Table 8.3) can be entered into this equation to yield the following:

$$d2_s = [(31 - 79.855)^2 + (5 - 4.3086)^2 + (24.4 - 13.575)^2 + (30.5 - 27.6383)^2 + (6104 - 5908.132)^2 + (72 - 472.5353)^2 + (144 - 147.5911)^2 + (144 - 392.7546)^2]^{1/2}$$

$$= 513.0281$$

$$d2_{ns} = [(31 - 97.1958)^2 + (5 - 4.4515)^2 + (24.4 - 11.263)^2 + (30.5 - 24.705)^2 + (6104 - 5731.842)^2 + (72 - 916.4705)^2 + (144 - 365.1802)^2 + (144 - 794.7305)^2]^{1/2}$$

$$= 1152.6467$$

Since $d2_s < d2_{ns}$, 5LA5364 is closer to the site group mean values in the measurement space and is assigned to the site-present class.

In actual practice the data should be standardized so that each variable (dimension) contributes equally to the calculations. In cases where the variances for each variable for each class are equal and where the variables are uncorrelated, this algorithm minimizes classification error (Schowengerdt 1983:54). Even when these special conditions do not arise, studies have shown that the accuracy of the minimum distance classifier is comparable to that of the maximum likelihood method (Hixon et al. 1980). The minimum distance classifier may be calculated using the Statistical Analysis System PROC NEIGHBOR program (SAS Institute 1982).

Figure 8.8a maps the results of applying the minimum distance ($d2$) classifier to the 19,000 locations in the test study area described above. The shaded areas are those portions of the region that were classified as being closer to the site class mean values in the measurement space than to the nonsite values. This "site-similar" region compares favorably with the site subspace delineated by the statistically derived maximum likelihood classifier (for comparative purposes, the site subspace defined by maximum likelihood as all locations with conditional site probabilities > 0.5 is portrayed in Figure 8.8d). Similarly, the classification accuracy and gain statistic of the results of applying the $d2$ classifier to the sample data of 1423 locations (Table 8.4) compare favorably with those of the multivariate data analysis procedures of the previous section, if the relative difficulties of the two types of procedures are taken into account.

A slightly different distance measure, termed *city block* distance (Schowengerdt 1983:51), is merely the sum of absolute distances from the ith location to the class mean values:

$$d1_k = |x_{1i} - \mu_{1k}| + |x_{2i} - \mu_{2k}| + \ldots + |x_{ji} - \mu_{jk}|$$

This distance measure is somewhat easier to employ than the Euclidean distance measure because it requires fewer calculations. When mapped across the study region (Figure 8.8b) this decision rule yields results almost identical to those of the $d2$ classifier (Figure 8.8a). The classification accuracy and gain statistic for the application of the $d1$ rule to the 1423 sample locations are given in Table 8.4, and these, too, are almost identical to those for the $d2$ results. Despite these similarities, the $d2$ rule is more commonly used because it is more interpretable and tends to perform somewhat better than the $d1$ rule.

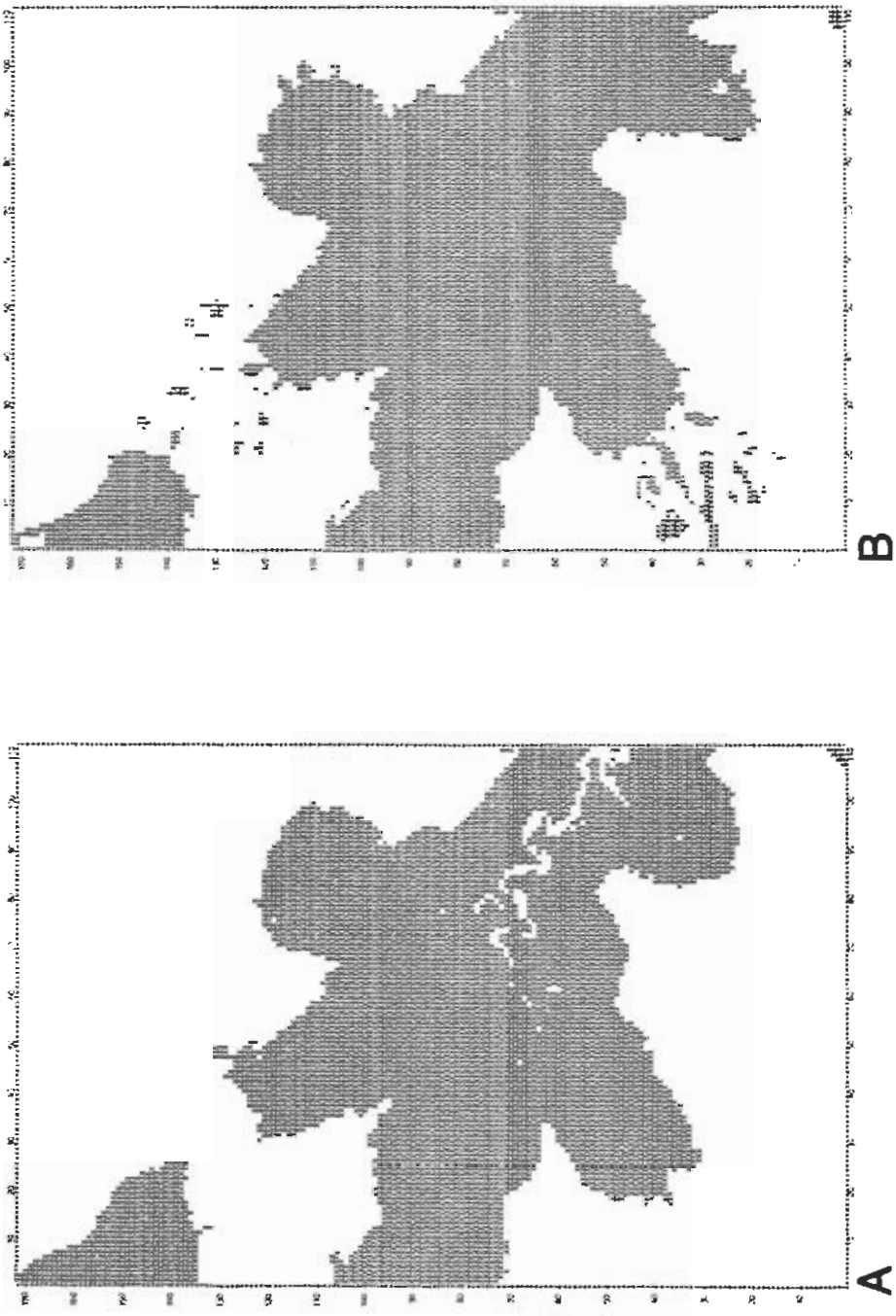
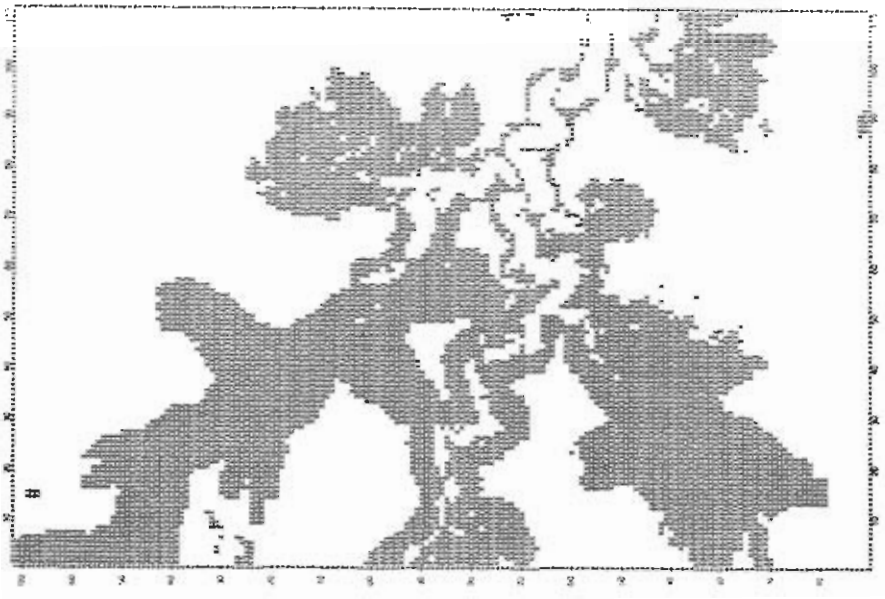
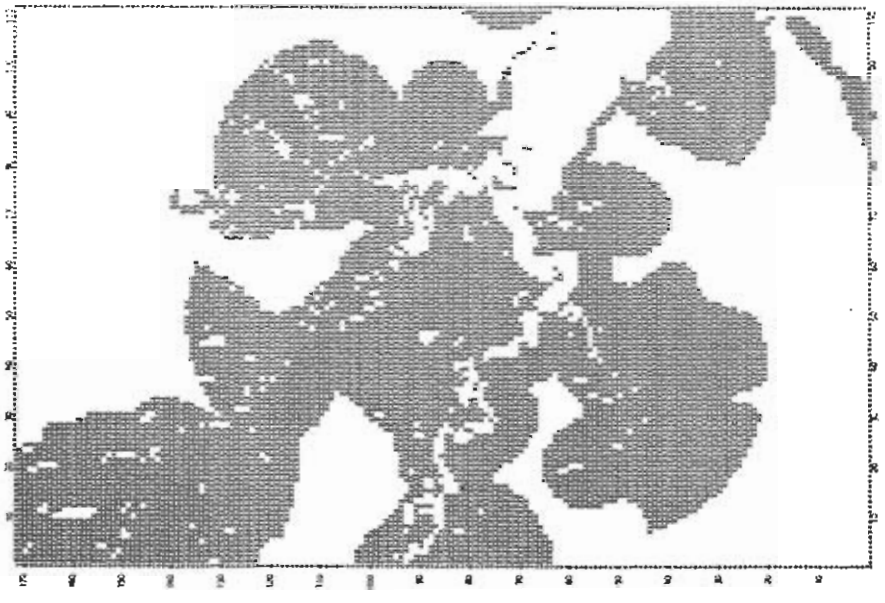


Figure 8.8. Simple mathematical models based on characteristics of locations (eight environmental variables): (A) minimum (Euclidean) distance (*d*2), (B) minimum (city block) distance (*d*1), (C) level-slice classifier (±1.75 *s.d.*), (D) maximum likelihood classifier at $p \geq 0.50$.

(continued)



D



C

Figure 8.8. Continued.

Level Slice Classifier

This algorithm is sometimes called the *parallelepiped classifier* (Moik 1980:271–272). It establishes decision boundaries that are parallel to *each* axis in the measurement space (hence the term *level slice*) by forming a hyperrectangle or parallelepiped about the class(es) of interest; a hypothetical level-slice decision boundary in a measurement space is portrayed in Figure 8.5b. In image processing, minimum rectangles are usually fitted around class boundaries derived through maximum likelihood estimation. The best known archaeological application of this technique is the polythetic choice model of prehistoric settlement developed for the Reese River Valley of Nevada (Williams et al. 1973). This polythetic choice model used arbitrary cutpoints (level slices) on each of seven environmental variable axes. In this case, however, a location was classified to the site-present group (no actual nonsites were used) if any five of the seven values measured at a location were below the threshold levels. The level slice classifier can be defined for any variable as

$$(\bar{x} - t) \leq x_i \leq (\bar{x} + t)$$

where x_i is the value of a variable at the ith location, $\bar{x}$ is the estimated mean value for the class of interest, and t is a threshold or cutpoint value (see Moik 1980:273).

To illustrate application of this method with our example data, measurements from 5LA5364 and the mean and standard deviation data from the site group sample were used to produce the results shown in Table 8.4. The threshold value, t, is arbitrarily set at ± 1.75 standard deviation of the site group mean. Note that any threshold value may be selected and that this choice will directly affect subsequent performance; in the present case, several values of t were examined before the ± 1.75 s.d. value was selected because of its relatively good performance. Inserting the relevant data for each variable, we find that the following relations hold:

$\bar{x}$		t s.d.		*5LA5364*		$\bar{x}$		t s.d.
79.8550	−	(1.75)52.7906	≤	31	≤	79.8550	+	(1.75)52.7906
4.3086	−	(1.75)4.1598	≤	5	≤	4.3086	+	(1.75)4.1598
13.5750	−	(1.75)10.4165	≤	24.4	≤	13.5750	+	(1.75)10.4165
27.6383	−	(1.75)18.3213	≤	30.5	≤	27.6383	+	(1.75)18.3213
5908.1320	−	(1.75)460.7342	≤	6104	≤	5908.1320	+	(1.75)460.7342
472.5353	−	(1.75)637.6298	≤	72	≤	472.5353	+	(1.75)637.6298
147.5911	−	(1.75)165.4219	≤	144	≤	147.5911	+	(1.75)165.4219
392.7546	−	(1.75)465.3968	≤	144	≤	392.7546	+	(1.75)465.3968

Thus, 5LA5364 is classified to the site group.

Application of this procedure to all 1423 locations in the sample yields the accuracy rates and gain statistic given in Table 8.4 for the ± 1.75 s.d. threshold value. When the level slice is applied to each of the 19,000 locations of the test study region, the resulting mapped subset (Figure 8.8c) is very much like those resulting from the application of minimum distance (Figure 8.8a and b) and maximum likelihood (Figure 8.8d) classifiers. As with all of the above techniques, it is quite

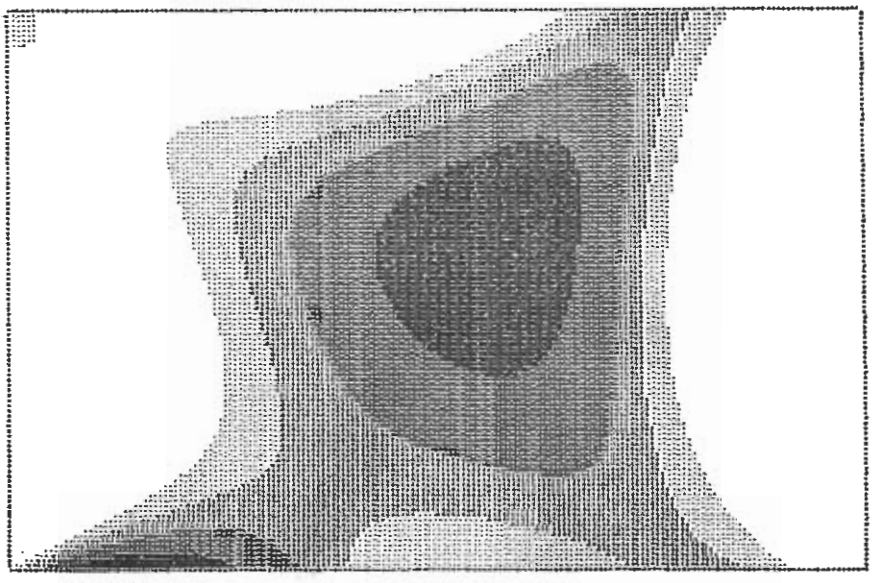
easy to alter results in either direction simply by changing a cutoff point or threshold value.

COMBINING MODELS FOR LOCATIONAL CHARACTERISTICS AND MODELS FOR LOCATION ONLY

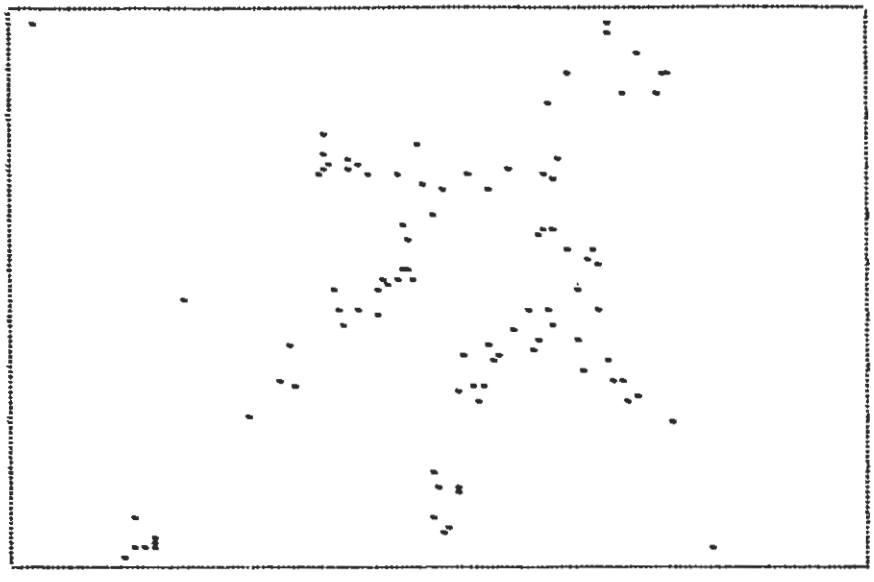
A fundamental dichotomy in types of archaeological locational modeling approaches was established early in this chapter. Models were classified as those based only on locational data (spatial x,y coordinates) or those based on characteristics of the locations, such as environmental information. A fourth-order polynomial logistic regression model was presented as a model based only on locational data in the section "Approaches Based on Trends in Location Only." The preceding sections have illustrated several approaches for models based on the environmental characteristics of locations. Since both approaches to modeling provide information—and generally independent information—summarizing where archaeological sites are located, it would seem a logical step to combine these approaches in order to enhance our ability to model prehistoric site distributions.

To conduct this experiment the Colorado plains study region is used again. Unlike the analysis in the previous section, which utilized 1423 site and nonsite locations from the entire 575 km² study region, the present analyses make use only of the samples of 95 site-present and 54 site-absent locations from the 5.5 by 8.5 km study portion of the larger region. (The smaller region has been used in Figures 8.4, 8.7, and 8.8 to portray various model mappings.) This smaller region is used here for two reasons. First, the logistic trend-surface technique for location only is best suited for modeling the reduced complexity of a smaller region (and such a model has already been established for the present region in Figure 8.4). Second, the environmentally based models of the previous section were based on locational patterns from a collection of sites from a huge region; therefore, these models averaged the locational pattern of all the sites and nonsites from the wider region. It is germane to illustrate the power of the environmentally based approach by fitting such a model to a relatively small region, which must contain a lower degree of environmental variability than the larger study area and therefore offer the potential of a tighter fit of the model to the data.

In order to facilitate comparison of these modeling approaches, the locations of the 95 site-present cells (out of nearly 19,000 cells) in the study region are shown in Figure 8.9a. The logistic trend-surface model derived through use of the fourth-order powers of the spatial (x,y) coordinates of the 149 site and nonsite locations is shown again in Figure 8.9b. The classification accuracy of this *location-only* model for these data is given in Table 8.5. The pseudo-R^2 goodness-of-fit statistic (defined earlier) for this model is $R_p^2 = 0.5318$, and the gain statistic is estimated as $1 - 31.5/82.1 = 0.616$.



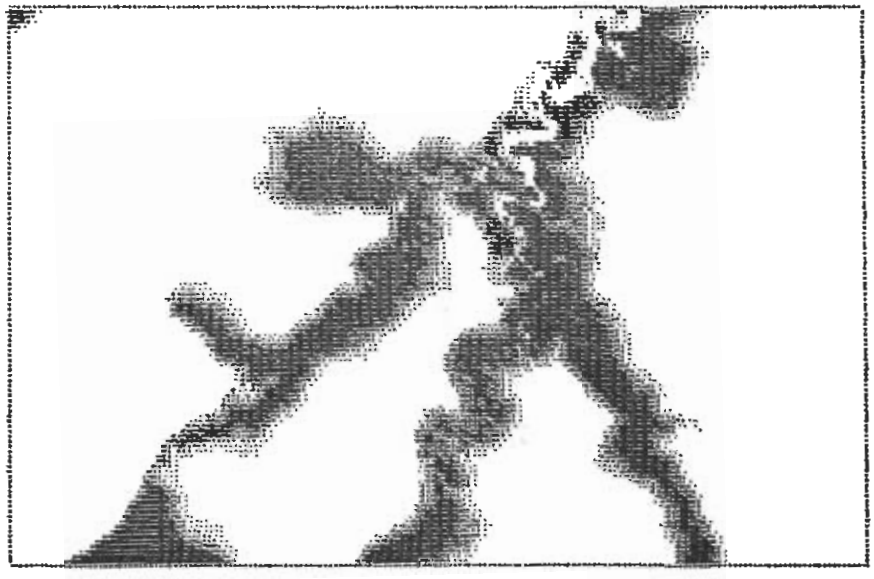
B



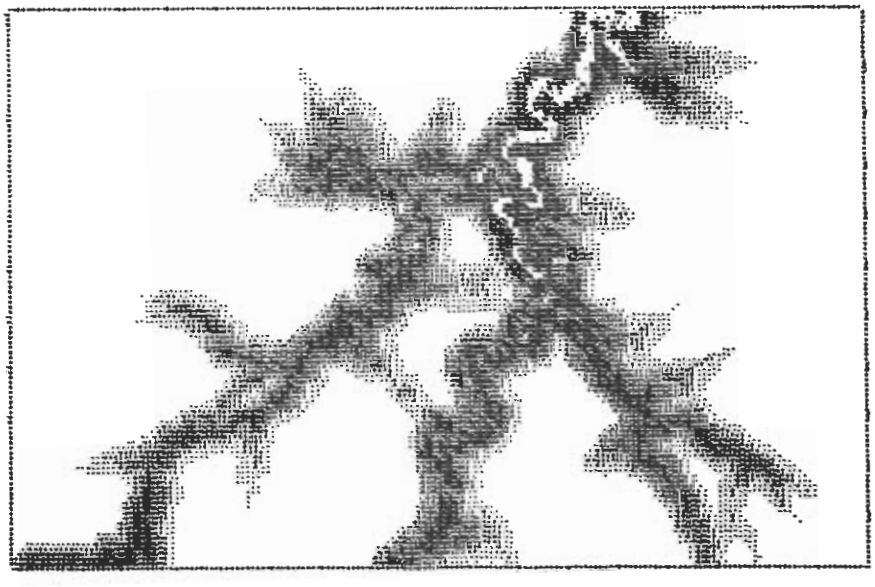
A

Figure 8.9. Combination of location-only and locational characteristics models: (A) actual site distribution, (B) fourth-order logistic trend surface based on x,y site coordinates, (C) logistic regression model using eight environmental predictors, (D) combination of locational (B) and environmental (C) characteristics models.

(continued)



D



C

Figure 8.9. Continued.

TABLE 8.5.

Comparison of classification performances of site location models for (a) locational coordinates, (b) locational characteristics, and (c) locational coordinates and characteristics for a portion of the Colorado plains study area (row percents given in parentheses)

Locational Coordinates Predictions			*Locational Characteristics Predictions*			*Locational Coordinates and Characteristics Predictions*		
	Site	*Nonsite*		*Site*	*Nonsite*		*Site*	*Nonsite*
Actual	$p \geq 0.5$	$p < 0.5$	Actual	$p \geq 0.5$	$p < 0.5$	Actual	$p \geq 0.5$	$p < 0.5$
Site	78 (82.1)	17 (17.9)	Site	87 (91.6)	8 (8.4)	Site	90 (94.7)	5 (5.3)
Nonsite	17 (31.5)	37 (68.5)	Nonsite	9 (16.7)	45 (83.3)	Nonsite	7 (13.0)	47 (87.0)
$R_p^2 = 0.5318$			$R_p^2 = 0.7156$			$R_p^2 = 0.8081$		
gain = 0.616			gain = 0.818			gain = 0.863		

A logistic regression model for environmental *locational characteristics* was also fitted to the site-present and site-absent data in the 5.5 by 8.5 km study region using the eight variables described in the preceding sections (Figure 8.9c). Since this model is based only on the data patterns of the 95 sites and 54 nonsites in the smaller study area and not on all of the 1423 site and nonsite locations of the larger region, the resulting model provides a much tighter fit to the site data than the previous logistic regression model (Figure 8.7c). The classification accuracy statistics for this locational characteristics model are given in Table 8.5; $R_p^2 = 0.7156$ and gain equals 0.818. The fact that environmental characteristics of locations provide more information than simple locational coordinates is amply illustrated by comparison of the resultant models (compare Figures 8.9b and 8.9c).

Finally, a model was developed that combined *both* positional data and information about locational characteristics. This was accomplished by utilizing the 14 polynomial terms of the location-only model and the eight environmental terms of the locational characteristics model simultaneously in a single logistic regression model (Figure 8.9d). The results of this model exhibit characteristics of both the trend-surface and the environmental models, as indicated by the mappings (Figure 8.9b-d). By incorporating both sources of information, classification accutacy is increased (Table 8.5), as suggested by the higher goodness-of-fit ($R_p^2 = 0.8081$) gain (0.863) statistic, and model mapping.

MODELING INDIVIDUAL SITE TYPES

The methods discussed in the previous sections may be applied not only to questions of site presence and absence but also to modeling multiple site types within a region, as has been noted. These might be functional site types or site

types representing different chronological periods, for example. The major problem in developing locational models for individual site types lies not in the methodological difficulties of developing the models but in the definition and operationalization of meaningful site-type categories and in acquiring sufficiently large samples of the types for analysis. These problems were discussed in greater detail in the introduction to this chapter.

When dealing with probabilistic locational models of archaeological phenomena, it is often desirable that the individual class p-values (probabilities of class assignment) for all of the classes under consideration sum to 1.0 for any location to which the model(s) are applied. A number of standard and not-so-standard procedures exist that allow one to constrain estimated probabilities from multiple groups (e.g., site-type groups) to sum to 1.0 at a location.

The simplest model is one that assumes for any location (a small land parcel, such as an acre) a limited and finite number of possible outcomes and then estimates the probability of any given alternative (e.g., by means of some of the modeling procedures discussed above). In an archaeological context the alternatives that may occur at a location include an alternative for each possible site type (including isolates or other remains) and the alternative of no site (no archaeological remains of any kind), and

$$p(ns) + p(st_1) + p(st_2) + \ldots + p(st_n) = 1$$

where ns is the site-absent alternative and the st_i refer to the individual site types. This model assumes that all possible site types have been specified, but this difficulty may be circumvented simply by defining a type called "other." This kind of model is assumed by many packaged computer programs for statistical analysis, including those for multigroup discriminant analysis (e.g., PROC CANDISC; SAS Institute 1982).

An alternative model that perhaps offers a number of advantages, given our limited knowledge of the past and the normal difficulties of dealing with the archaeological record in terms of defining site types, is a hierarchical model that first assumes only two possible outcomes at any given location (again, a small land parcel, such as an acre). One outcome is that some evidence of human activity (ha) occurs at the location (i.e., some kind of site or cultural manifestation will be found there); the other is that no evidence of activity (ns) occurs at the location, and

$$p(ns) + p(ha) = 1$$

Outcomes indicating specific kinds of activities, archaeologically represented by functional types of sites or remains, are then conditional on the outcome that evidence of human activity is indicated (Wrigley 1982), and

$$p(st_1) + p(st_2) + \ldots + p(st_n) = p(ha)$$

The term $p(ha)$ refers to a human activity space within which all activity in a region is conducted (see also the introduction to this chapter). This space is represented archaeologically by all material culture remains, including settlements, sites of specific function, and isolated occurrences (see Kvamme 1985a for a more detailed

discussion of this concept). Although some researchers might argue that in certain regions past activity occurred everywhere, the concept of importance here is that of *activity densities*: in any region certain locations may have been more favorable for activity *of any kind* than others (e.g., locations with level ground surfaces). In fact, in mountainous regions or in regions containing significant acreage of swamplands, for example, the human activity space can be restricted to a major extent. Site location models that lump sites of all types into a single analytical group—whether because of an inability to form meaningful types from the available evidence or as part of a preplanned research tactic—are simply developing models of the human activity space. Such models should, in principle, demonstrate less-pronounced patterning than models for specific types of sites since the former incorporate many types of sites with varied locational requirements. Nevertheless, strong and predictable patterning can sometimes be achieved (e.g., Figure 8.9d portrays a remarkably strong model for all open-air lithic scatters in the region even though these scatters undoubtedly represent a variety of functional site types). This hierarchical scheme has the advantage that locational models for specific site types may be incorporated as well (e.g., at a later time). The site-type models, however, are conditional not only on the environmental and other measurements upon which they are based but also on $p(ha)$.

The following illustration of locational models for multiple site types uses data from a study of German Mesolithic sites by Kvamme and Jochim (1988). This study was used as an example of model building with existing data in Chapter 7. As noted there, the data available for the 170 known sites in the region were extremely limited, making it difficult to distinguish site types with specific functions. The amateur collectors who discovered them, however, reported a number of sites that contained relatively abundant remains. Of the 170 sites, 39 could be categorized in a "settlement" group and 74 were classed in a "small-sites" group (the remaining sites were unclassified or represented "isolates" and are not used here). Although the validity of these site types is questionable, we can assume for the purposes of this discussion that the types *are* valid and use these data to illustrate the simultaneous modeling of multiple site-type groups within a single region.

An advantage of using these Mesolithic data as an example of site-type modeling is that a computerized GIS has been established for the entire 940 km^2 study region (see Chapter 10 for a discussion of geographic information systems). The GIS contains a gridded representation of the study region comprising approximately 84,000 cells measuring 100 m on a side. Within each grid cell values were estimated for elevation, slope, aspect, local relief, a measure of view, a measure of shelter, horizontal distance to nearest water, vertical distance to nearest water, and horizontal distance to nearest third-order stream. Additionally, grid cells that contain Mesolithic sites were denoted and information on the site types present in the cell was encoded. The mathematical details of the site-rype models are presented in Kvamme and Jochim (1988) and Kvamme (1986); for the purposes of this discussion, "pictures" of each model will be used to indicate the results of site-type modeling. These mappings were accomplished by applying the models across the

entire study area (i.e., in each of the 84,000 cells) and using computer cartographic techniques to display the results.

The site location models were developed by contrasting the 39 settlement locations (cells) and the 74 small-site locations against a representative sample of 3201 locations taken from the background environment at large, rather than from a group of known site-absent locations (because the latter information was unavailable). The site types could be contrasted with the background environment because Mesolithic sites of any type could be argued to be an extremely rare phenomenon, causing the background environment to constitute a reasonably distinct class for analysis purposes (in other words, in the sample of background locations only a very small percentage of the selected locations could be expected to contain as yet undiscovered Mesolithic sites by chance). The classifier used was logistic regression (discussed above), and all nine variables listed above were incorporated in the models.

The study region consists of high ridges and plateaus overlooking a number of river valleys and plainslike areas (Figure 7.4, this volume). The analysis suggested that sites of the settlement class tend to be located at lower elevations, on less sloping ground, in regions of less local relief that were more sheltered (i.e., less likely to be on hilltops and more likely to be on valley bottoms), and in places with lower values for the overall view measure than the small-site class. Additionally, sites of the settlement class tend to lie closer to relatively secure water sources, including major drainages (third-order streams), although they did exhibit a slight orientation toward location at greater distances from nearest drainages when compared with the small-site class.

These findings are largely borne out by the site type locational models when they are mapped over the entire region through a GIS (Figure 8.10). The map of the model for the settlement class (Figure 8.10a) shows a locational pattern (the darker regions) emphasizing areas along a major drainage near the southwest border of the region and in a plainslike low-elevation area in the far western portion of the region. The locational pattern mapped for the small-site class (Figure 8.10b) does not show an emphasis on these areas. More important, however, the locational models indicate that the settlement class is much more highly patterned in terms of location than the small-site class, as indicated by the relative sizes of the dark areas in the two maps. In other words, the settlement class tends to exhibit a tighter and more restrictive locational pattern while the small-site class pattern seems to be more variable.

Elsewhere I have shown with regional survey data from the western United States that large sites or settlements do indeed tend to be more patterned in location than smaller sites, which as a group are functionally more variable (Kvamme 1985a). In that study, it appeared that the greater functional variability within a small-site class led to greater locational variability (presumably owing to different locational requirements of individual functional site types that were pooled within the small-site class) and that these sources of variability caused the less-pronounced locational patterning of these sites. The large-site or settlement

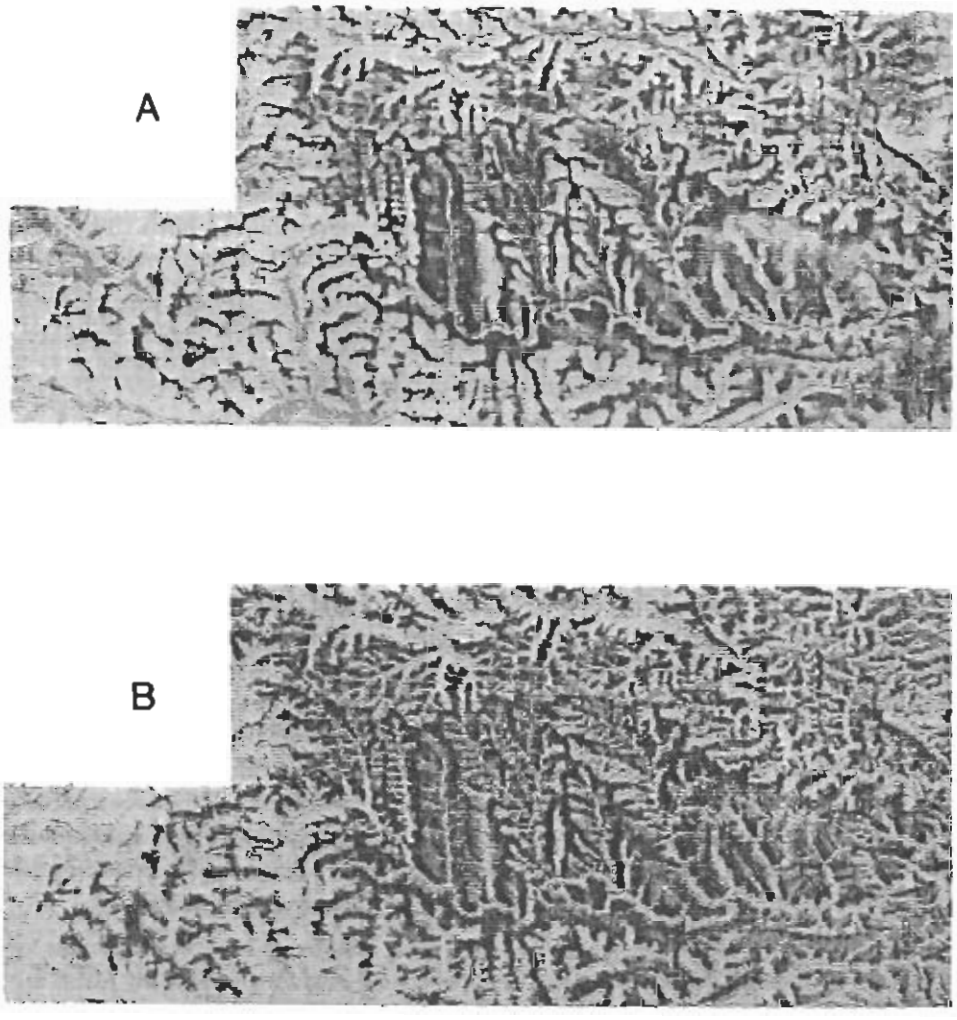


Figure 8.10. Mesolithic site-type locational models mapped through a GIS: (A) settlement class, (B) small-site class.

class, on the other hand, possibly consisted of sites representing a more similar range of activities (e.g., extended occupation) with similar and thus tighter locational requirements. Although the integrity of the site classes and sample in the Mesolithic study may be questioned, the patterning discerned does resemble the large-site/small-site patterning discussed here.

INTERPRETATION AND EXPLANATION OF DATA PATTERNS

The foregoing sections have presented a number of quantitative data analysis techniques that are relevant in the development of archaeological locational models. As scientists, archaeologists are also interested in *(a)* methodological rigor and *(b)* explanation. The methods presented offer great potential for both. The most obvious benefit obtained from use of quantitative methods of data analysis is that research findings can be obtained with greater objectivity. Additionally, results tend to be more easily replicated by other investigators: another researcher can duplicate an experiment or analysis using identical procedures and the same or even similar data, allowing independent verification of findings.

Quantitative analysis procedures yield other benefits that may be less obvious. In traditional archaeological research strategies the researcher only has access to *(a)* the raw phenomena (objects, entities, individuals) under investigation; *(b)* data (observations or measurements) pertaining to those phenomena; and *(c)* relationships subjectively perceived between and among the data or phenomena. The researcher who has knowledge of and access to empirical data analysis methods, on the other hand, can greatly augment these most fundamental capabilities because these procedures yield additional information in the form of *(d)* descriptive and summary statistics, which describe and generalize tendencies and patterns in the data and make relationships explicit; *(e)* complex data models, which portray the raw data in different ways, often illustrating or summarizing the essence of multiple empirical patterns; and *(f)* unforeseen (multivariate) relationships between classes of phenomena. Thus, the practicing scientist who makes use of empirical analysis procedures can greatly increase his or her abilities to postulate a pattern among the basic facts of the discipline, an important basis for theory formulation.

Quantitative methods of analysis are also beneficial in other domains of research. In classic deductive research approaches, certain predictions often are made based on the premises of the initial hypotheses. In the hard sciences these predictions usually rest on mathematical deductions or established physical laws such that the predictions *must* mathematically (or by law) follow from the hypotheses. In archaeology, which lacks a base of laws or theory, our bridging arguments that lead to predictions, as Thomas (1979) notes, are "seat-of-the-pants" kinds of statements. Well-established relationships of a statistical kind might be used here as an alternative or supplement to such arguments when predictions are formulated from theory. Finally, the methods of statistical hypothesis testing are particularly well suited as a means of verifying (or refuting) hypotheses. A myriad of testing and validation procedures exists for virtually any type of problem context. Hence, the quantitative investigator is armed with more tools and capabilities for conducting basic research and for potentially interpreting and explaining archaeological phenomena.

Previous sections generally presented only the basic statistical facts because their goal was to describe the procedures used in modeling. In this section, the

interpretation and explanation of these facts are briefly considered. The Glade Park descriptive statistics in Table 8.1 suggested a number of empirical tendencies. For example, the sites exhibited tendencies to be located in proximity to permanent and nearest water sources (when contrasted to the background nonsites) and also tended to be located with good views of surrounding terrain and close to points of vantage. The Glade Park sites were distributed with a north-facing preference, on level ground, and on high points or mesa edges in regions of limited local relief. Traditionally, explanation of empirical patterns such as these have generally assumed human selectivity (e.g., Findlow 1980; Green 1973; Kvamme 1985a; Lafferty 1981; Parker 1985; Roper 1979b). For example, using the above evidence we might argue that sites tend to be close to water because the Glade Park region was arid, forcing people to carry out most of their activities in areas near water courses. The aridity argument might also explain the strong tendency for north-facing aspects since these locations would tend to increase shelter from sunshine during the hot summer if, indeed, the sites were occupied during the summer. Locational tendencies toward good views, vantages, and high points might be interpreted as resulting from the need to watch for game, since the inhabitants of Glade Park were primarily hunter-gatherers. Such arguments, although plausible, need to be substantiated with additional evidence before they are taken seriously. Alternative explanations might also be possible.

The actual equations of the empirical models based on characteristics of locations also have interpretive potential. The discriminant analysis discussed in the "Application Comparison" section yielded the following model:

$$D_i = -5.7058 - 0.0047\,(\text{aspect}) - 0.08\,(\text{slope}) + 0.0152\,(\text{relief, 100 m}) - 0.005\,(\text{relief, 300 m}) + 0.001\,(\text{shelter index}) - 0.0002\,(\text{vanrage distance}) - 0.001\,(\text{distance to nearest drainage}) - 0.0008\,(\text{distance to second-order drainage})$$

Positive coefficients associated with a variable suggest that high values of the variable are related to site presence, while negative coefficients suggest that low values of the variables are related to site presence. Hence, this model indicates that high values of relief (within 100 m) and the shelter index and low values of aspect (i.e., north-facing), slope, relief (within 300 m), vantage distance, and distance to nearest and second-order drainages are suggested by the data to be related to the site-present locations.

It is possible to go beyond this level of interpretation when the independent variables are measured in the same units and are uncorrelated. One way to acquire variables measured in the same units is to standardize the data. Parker (1985) utilizes this tactic to interpret logistic regression site location models in Arkansas. In the above nonstandardized discriminant analysis model, several variables are measured in the same units. We might compare the absolute values of the associated coefficients of these variables to assess the relative importance of the variables. For example, the distance variables are all measured in meters; if we compare them we find that the data suggest that distance to nearest vantage (with an absolute coefficient of 0.002) has about one-fourth as much influence as distance to second-order drainage (0.0008) and one-fifth as much influence as distance to nearest

drainage (0.001). Distance to nearest drainage is slightly more important to site-location placement than distance to second-order drainage. In the present case, however, these variables are positively correlated, and such interpretation should be made with some caution.

One way to remedy the correlation problem is to perform a principal components analysis (described earlier) on the variables. This procedure yields linear combinations of the correlated variables, and a model can then be built that uses the components rather than the raw variables as predictors (Schowengerdt 1983:159–167). Since the resultant components will be uncorrelated, model interpretation can be facilitated in this manner. Although this approach has certain merits, it often is the case that interpretation of the components themselves is quite difficult.

Explanation of any facts pertaining to archaeological distributions, whether raw facts or higher-order statistical generalizations, may take a number of forms. A good approach might be to treat each possible explanation as an alternative hypothesis. Possible alternative hypotheses for an observed relationship between archaeological sites in a region and some environmental feature might include *(a)* human selectivity, *(b)* geologic processes, *(c)* vegetation patterns, and *(d)* sampling biases.

To illustrate this multiple-hypothesis approach, recall that the models discussed in the section on "Application Comparison of Quantitative Locational Models" all indicate that the locations of open-air lithic scatters tend to occur in close proximity to second-order (or greater) drainages (Figure 8.7). In Chapter 10, a histogram of this variable measured at all 230,000 land parcels (50 by 50 m units) in the study region is compared with a histogram of the same variable measured only at the nearly 600 parcels with sites in the area (Figure 10.11). These histograms clearly support the suggested pattern; for example, half of the sites occur within 150 m of drainages of these ranks, while only 17 percent of the study region lies within this distance of such drainages.

The explanation of this pattern that probably comes to mind first is that of human selectivity: the prehistoric inhabitants purposefully placed their sites in proximity to relatively secure sources of water in order to obtain water more easily. Various sources of ethnographic evidence and the aridity of the southern Colorado plains could be argued to be supportive of this hypothesis. An obvious and related alternative hypothesis is that the inhabitants tended to locate activity close to drainages not for the water but for some other related resource. For example, they might have been exploiting plants that tend to be found near water, or they might have chosen stream-associated locations in order to hunt a variety of game animals, such as bison, that might be drawn to water. This is a common case, where one variable (proximity to water) might be only a proxy for some other variable (availability of plant foods or prey animals) that actually was important. Supporting data for this competing hypothesis would be hard to obtain. Such data might include appropriate floral and faunal remains in a suitable archaeological association

from sites both close to and far from the drainages. In addition, the sample of sites would need to be large enough to yield statistically reliable conclusions.

A third hypothesis might be that the observed pattern is a result of geological processes or vegetation patterns that have buried or hidden sites located great distances from these drainages and exposed sites lying in proximity to the drainages. In the present case, a geomorphological study of the region (Schuldenrein 1983) found the reverse to be true; the primary areas of alluviation were along major drainages. Additionally, vegetation in this region (which affects site visibility) is densest along major drainages and very light or nearly absent far from drainages (Van Ness 1984).

Finally, a fourth explanation of the pattern might be that it is the result of sample selection bias. Since a random sampling design was employed for site discovery (based on randomly placed transects), and assuming the trustworthiness of the survey crews and uniformity of their procedures, this hypothesis seems an unlikely candidate.

Certainly there are other alternative hypotheses for explaining the observed relationship. In this case, as in all cases involving hypothesis testing, the alternative for which the greatest amount of supporting evidence can be obtained should be advanced as the most likely explanation. It is also quite possible that several of the hypotheses could be true.

ASSESSING MODEL PERFORMANCE

In previous sections initial or "apparent" accuracy rates were presented for several models. Apparent accuracy rates were obtained by applying a model to the *same* data used to generate the model. This practice, as noted in those discussions, tends to give an inflated view of true model accuracy and underrepresent true model error rates. The purpose of this section is to examine methods that can yield truer indications of actual model performance and to offer statistical significance tests of model performance. It is emphasized that regardless of how a model is developed—from theoretical expectations or from empirical data—most of the following methods for testing apply. These methods should be used to validate the performance of *any* model prior to its application to management or research problems.

Adjustable Accuracy Rates

Site location models discussed in previous sections were designated as having classified a percentage of sites correctly and a percentage of nonsites correctly. Some of the models classified only about 70 percent of the sites correctly (and some had a lower rate than this), which might not be very useful from a practical standpoint;

the 70 percent correct figure means that 30 percent or more of the sites were incorrectly classified. This is fairly costly given the nature of the resource and our goal of developing models that have some potential for real-world application. One way to resolve this problem might be to obtain better data or to make operational new variables that would yield stronger models, but either solution could entail additional cost and effort. Even if such models were developed, some sites and some nonsites will always be incorrectly classified by a model, and the accuracy rates might be less than desirable or be unacceptable for practical applications.

A solution to this problem is to accept a trade-off, to exchange increased accuracy in classifying sites for decreased accuracy in classifying nonsites since it costs us less to call a nonsite a site compared with the reverse (using the terminology introduced in Chapter 3, we decrease gross error by increasing wasteful error). The decision rule used for examining the initial apparent accuracy rates of all previous models (e.g., Tables 8.2, 8.4, and 8.5) was a maximum-likelihood rule; locations were assigned to the class (site or nonsite) to which they were most similar. For several of the models this amounted to assigning a location to the site class based on a cutoff point of $p = 0.5$. In order to trade nonsite classification accuracy for increased site accuracy, we need only change this p-value to a lower cutoff—for example, to $p = 0.25$. In terms of the measurement space (Figure 8.5b), this change moves the decision boundary upward, causing more of the sites to be correctly classified (but causing more nonsites to be incorrectly classified). The logical extreme for this tactic would be to choose a cutoff of $p = 0.0$, which would cause the entire measurement space to be classified to the site group (but this would create the absolutely accurate but useless predictor of site locations described in an earlier section).

We can use the Glade Park nine-variable site location model presented in the "Example Analysis" section as an illustration of this procedure. This model is based on a sample of 157 known site and 157 known nonsite locations, obtained through a cluster sample of 38 quarter-section quadrats. The Glade Park model was applied to estimate site-group p-values for each of these 314 locations based on their nine environmental measurements. Histograms of these p-values are given in Figure 8.11a. If we use the traditional cutoff point ($p = 0.5$), 70.1 percent of the sites fall above this point in the site histogram, while 66.2 percent of nonsites fall below this cutoff in the nonsite histogram. (It is this process that yields the accuracy rate predictions given in the two-by-two matrix in Table 8.2b.) If we were to use a lower cutoff, it is readily apparent from Figure 8.11a that more sites would be classified correctly and more nonsites incorrectly. This effect is summarized in Table 8.6 using cutoff p-values of 0.0, 0.1, 0.2, . . ., 0.9, 1.0, and it is graphed in Figure 8.11b. Of course, when cutoff probabilities of $p = 0.0$ and $p = 1.0$ are used, *every* location is classified either as a site or as a nonsite, respectively, and we have a zero-gain predictive model. On the other hand, at a cutoff of $p = 0.2$, 96.2 percent of the sites and 26.1 percent of the nonsites are correctly classified; at $p = 0.4$, 82.8 percent of the sites and 52.9 percent of the nonsites are correctly assigned, etc.

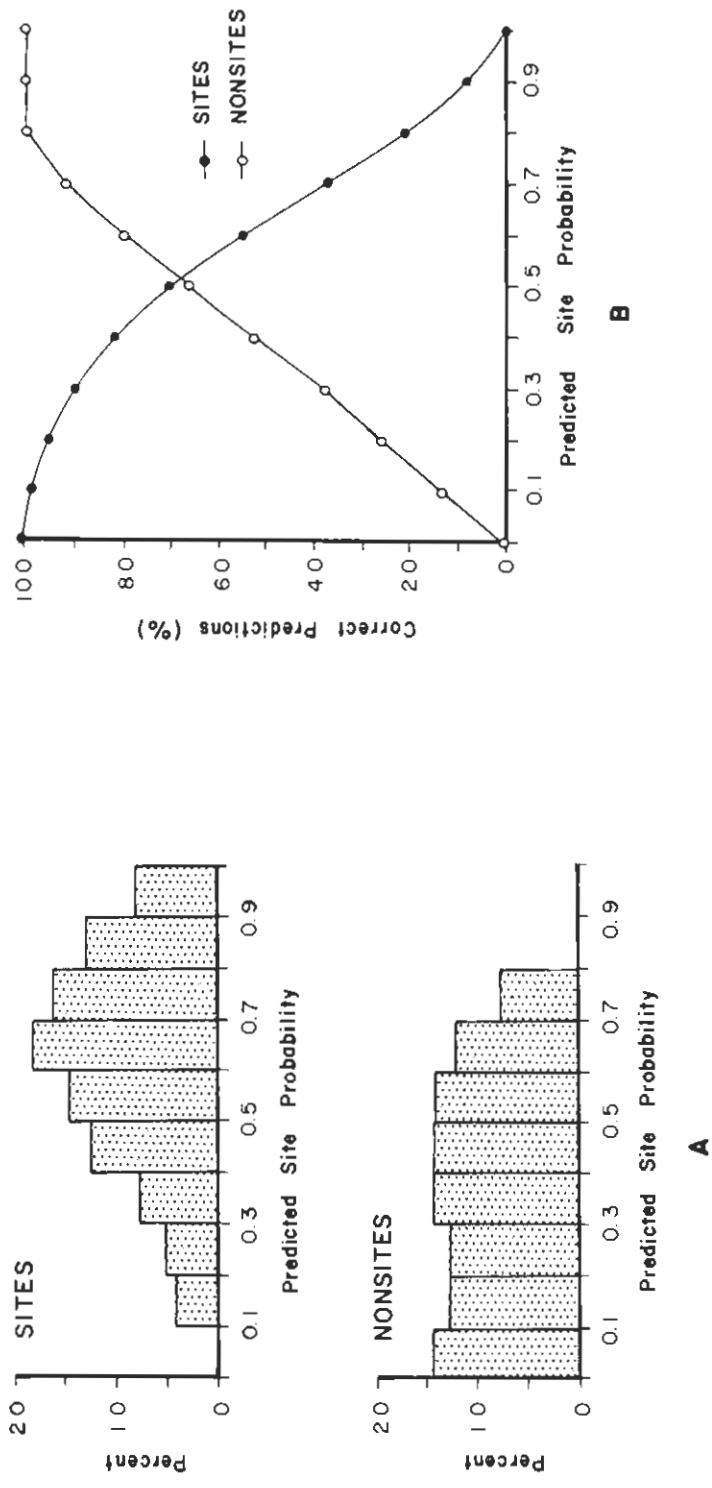


Figure 8.11. Estimated site-class p-values for the initial nine-variable Glade Park site location model: (A) histograms of the p-values for known sites and nonsites, (B) cumulative percent correct predictions for model sites and nonsites plotted for all model cutoff points.

Although application of a model to the same data used to build the model can yield inflated performance indications, the researcher can use these data, together with the adjustable rate method, to design a model that performs at an approximate level of accuracy. For example, using the information given in Table 8.6 and graphed in Figure 8.11b, the researcher can contrive a model that apparently predicts approximately 90 percent of the sites correctly by selecting a cutoff value of $p = 0.3$. At this cutoff point approximately 38 percent of the nonsites would also be correctly assigned. (Of course, these figures are undoubtedly inflated to some extent because the Glade Park model has not yet been tested with independent data. The same procedures apply after testing, however, and it is shown below that very similar results are obtained.)

TABLE 8.6.

Illustration of changing cutoff *p*-values and their effects on site and nonsite classification accuracy using the nine-variable Glade Park model data

	Correct Predictions				*Incorrect Predictions*			
	Site		*Nonsite*		*Site*		*Nonsite*	
Cutoff Point	*Number*	*Percent*	*Number*	*Percent*	*Number*	*Percent*	*Number*	*Percent*
0.0	157	100.0	0	0.0	0	0.0	157	100.0
0.1	157	100.0	21	13.4	0	0.0	136	86.6
0.2	151	96.2	41	26.1	6	3.8	116	73.9
0.3	142	90.5	60	38 2	15	9.5	97	61.8
0.4	130	82.8	82	52.9	27	17.2	75	47.8
0.5	110	70.1	104	66.2	47	29.9	53	33.8
0.6	87	55.4	126	80.3	70	44.6	31	19.7
0.7	58	36.9	144	91.7	99	63.1	13	8.3
0.8	33	21.0	157	100.0	124	79.0	0	0.0
0.9	12	7.6	157	100.0	145	92.4	0	0.0
1.0	0	0.0	157	100.0	157	100.0	0	0.0

Clearly, the actual number of locations (e.g., small-area units, such as acres) in a region of study that are nonsites is usually far greater than the number that are sites—on the order of 100 nonsites for every site. (This is usually referred to as the a priori or *base rate* probability problem and will be discussed below.) Thus, our claim that 38 percent of the nonsites are correctly assigned by a model essentially means that nearly 38 percent of the area of the study region as a whole is unlikely to contain sites (in this example, if the 38 percent area were mapped, it would only contain about 10 percent of all sites). If the study area is extensive this could amount to a sizable area that is largely devoid of sites. Thus, another important function of the nonsite control group is to provide area estimates about model performance. In other words, the nonsites provide data concerning the estimated area of a model at a particular cutoff point when mapped. If the data in Table 8.6 are correct indications

of the Glade Park model's accuracy, we could claim that 90.5 percent of the sites should occur in only about 100 - 38.2 = 61.8 percent of the total Glade Park land area, which could be called a "high site-sensitivity zone," and that 9.5 percent of the sites should occur in the other 38.2 pereent of the land area, which constitutes a "low site-sensitivity zone." Computer mapping techniques shown in the previous sections and described in detail in Chapter 10 may be used to provide maps of these sensitivity areas.

This cutoff adjustment approach is not necessarily restricted to models that yield a 0–1 scale of estimated probabilities. The simple mathematical models discussed in previous sections can also be examined in this context. The performance of minimum distance models might be assessed by investigating accuracy rates at various cutoff ratios of distance to class centroids, for example. Similarly, performance statistics from a number of "slices" might be examined in a level-slice approach.

From the foregoing it should be apparent that the use of overall accuracy rates (i.e., the combined site and nonsite accuracy) to evaluate the performance of a site location model, a fairly common practice (e.g., Berry 1984), is not only misleading but inappropriate. To illustrate, suppose that a sample survey from a large region discovers 100 site locations. It is possible to obtain virtually any sample size of nonsites as a control group since it is not uncommon for 99 percent of many study regions to be classifiable as "site absent." Let us say that 9900 locations are chosen for the nonsite control group, for a total sample size of 10,000. (Although this may seem to be a ludicrously large number, such sample sizes are possible through use of computer data bases, as Chapter 10 will show.) If all 10,000 locations were arbitrarily classed as nonsites, an impressive overall accuracy rate of 99 percent would be achieved (100 × [total correct] / [total cases] = 100 × [9900 + 0]/[10,000]), but the resulting model would be useless. It is clear that performance must be judged by focusing on percent correct rates for sites and nonsites individually.

Model Validation Procedures

The nine-variable Glade Park model has already been used to illustrate performance adjustment; in this section it will be used to demonstrate several model validation techniques. We can assume that the apparent performance rate statistics given in Figure 8.11b and Table 8.6 are inflated, but to an unknown degree. In other words, when this model is applied to other locations in the study region, actual accuracy rates may be lower than those indicated in the figure and table. The inflated performance statistics result from a number of factors. Primary among these is that the same data were used to build the model and to estimate the percent correct prediction rate (Table 8.6). Since the Glade Park model is based on differences between site and nonsite locations in that specific sample, the statistical procedures capitalize on variation in that sample such that apparent accuracy rates are maximized (Swain 1978:163). Violations of statistical assumptions, such as the

independence assumption that results from spatial autocorrelation (see above and Chapter 5), further widen the difference between apparent and actual accuracy rates, particularly in this cluster-sampled context (Basu and Odell 1974; Campbell 1981; Tubbs and Coberly 1978).

Randomization procedures for assessing the upward classification bias that results when a model is applied to the same data with which it was generated have been developed by Frank et al. (1965). These procedures were applied by Berry (1984) in a paper that generally attempted to discredit certain cultural resource modeling approaches. The procedures of Frank et al. (1965) and the findings of Berry (1984) are both germane to this discussion. One randomization procedure requires generating random normal deviates as a "synthetic" validation sample, and then developing a model based on these random data. The resultant classification accuracy, when the model is applied to the synthetic data set, reflects upward bias attributable to the procedure itself, since the robust properties of many multivariate classification models can cause a better-than-chance fit even to random data. Berry (1984) points out that in two such simulations by Frank et al. (1965), which used the discriminant analysis model, average overall classification rates of 68.2 and 72.6 percent were achieved, which would seem to reflect poorly on discriminant analysis efforts in general, including those in archaeology. Berry does not mention, however, that one simulation used 19 predictor variables with 150 cases and the other used 25 variables with only 98 cases (Frank et al. 1965:256). The large number of variables relative to the number of cases is an example of what can be called *hyperfitting* of a model to the data. It is possible, through use of large numbers of predictor variables, to obtain very strong fits regardless of the degree of patterning in the data (using n-1 predictor variables in a discriminant analysis guarantees a perfect classification, for example). This property is demonstrated in Berry's Table 2. Using random data and 30 cases, Berry shows through simulation that four variables yield an overall correct rate of 53 percent (3 percent upward bias); 8 variables, 70 percent (20 percent upward bias); 12 variables, 77 percent (27 percent upward bias); and 20 variables, 93 percent (43 percent upward bias). The difficulty in real-world applications is to obtain a good fit with few variables relative to the number of cases. This randomization procedure seems useful, however, when the numbers of variables and cases are matched to those actually used to develop an archaeological model. The results could give an excellent indication of the size of upward bias an investigator might be facing.

The second randomization procedure for investigating upward bias described by Frank et al. (1965) utilizes the actual model data for the predictor variables. In this case, though, the true value of the dependent variable, class membership, is randomized and a classification model is produced based on the randomized *groups*. The advantage of this procedure is that the actual model data are used, allowing the upward bias result to pertain more closely to the model under investigation. Berry (1984:849) utilizes this technique, with 10 replications, to illustrate a mean randomized classification rate of 71.6 percent, suggesting that the apparent overall accuracy rate of 85 percent for an archaeological locational model developed for the Bureau of

Land Management is mostly attributable to upward bias. The particular model that Berry examined was based on six vatiables and 174 cases (Burgess et al. 1980). Berry achieved his result, however, by using the six variables *plus* 10 additional ones that were eliminated from consideration in the original study owing to the lack of significant findings of several univariate and multivariate tests made on these variables. Thus, Berry achieved a mean randomization rate of 71.6 percent by hyperfitting 16 variables to 174 cases. I reran this randomization procedure with 10 replications on the original six variables, which yielded overall classification rates ranging from 49.4 to 60.9 percent with a mean rate of 56.9 percent, an upward inflation of less than 2 percent (the rate expected by chance in this case is 55.6 owing to unequal class sample sizes; Berry 1984). These findings are more in line with the amount of upward bias one might expect when the number of predictor variables is small relative to the number of cases.

There are a number of ways to conduct independent tests of a site location model's performance. In an *independent test*, data that are different from the information used to build a model are used to test the model in order to eliminate the possibility of model capitalization on chance sampling variation. The strongest test of model performance would require an additional independent survey. Site location models could be applied to these independent data to derive unbiased estimates of model performance accuracy rates. (Ideally, the independent survey would be conducted by archaeologists different from those who collected the data used to construct the initial site location model.) In many cases it is difficult or impossible owing to cost constraints to conduct a second, independent survey. For this reason, a number of alternative procedures have been developed that attempt to provide independent testing information but do not require that a second survey be performed. Two of these procedures, which were introduced in Chapters 5 and 7, are *split sampling* and the *jackknife* method.

Split Sampling

Split sampling traditionally requires randomly splitting a sample of cases (sites and nonsites) in half, building a model with one half, and testing the model with the second, independent half (Mosteller and Tukey 1977:38; see Chapter 7). A problem with this method results from the use of cluster sampling. There is within-cluster spatial correlation between analysis locations so that sites and nonsites in one of the split groups may not necessarily represent completely independent information relative to sites and nonsites in the other group.

A better split-sampling technique for cluster-sampled data requires that the *clusters* be randomly split into two groups of equal size. The model is then built with data from one-half of the clusters and tested with the second half, which now can be argued to be independent of the first half. This approach was applied to the Glade Park analysis data. The 38 sampling quadrats were randomly split into two groups of 18 (two of the quadrats contained neither sites nor nonsites and are excluded here); models were then built using the same nine variables used in the previous Glade

Park analyses for each half, and their performance was assessed using the data from the other half. The classification accuracy curves for all possible model cutoffs are illustrated in Figure 8.12a, as are the average performances of the two models. The inflation in accuracy of the original model, which amounts to a few percentage points, can be seen when the performance curves in Figure 8.12a are compared with those in Figure 8.11b.

A drawback of the split-sampling approach is that only half of the available information is utilized in developing a model (since the other half must be reserved for testing), which is a waste of costly information. One approach that utilizes all of the available data is discussed in the next section.

Jackknife Methods

The jackknife method (Lachenbruch and Mickey 1968) was developed as a means of providing a less biased assessment of the performance of a classification model while allowing all information to be used in model construction. In the traditional jackknife, one of the n cases is temporarily discarded, and the remaining n-1 cases are used to build a classification model. The discarded case is then independently classified by the model. This procedure is repeated, eliminating each case in turn, to establish an independent test of model performance. Thus, unlike split sampling where half of the cases are normally discarded, the jackknife requires that only one case be left out at any one time, which allows retention of most of the information. An additional benefit of this procedure is that the n resulting models, each providing a slightly different result, can be combined into a single model to provide a better estimated or jackknifed model (Mosteller and Tukey 1977:152). A model derived from n models is usually superior to the traditional model based on n cases because each coefficient in the combined model is based on n estimated coefficients from the individual models. The BMDP discriminant analysis program 7M provides the jackknife as an option (Dixon et al. 1983).

In an archaeological site location modeling context, where some form of cluster sampling is normally applied, a modified jackknife procedure can be used. This is necessary because, as noted in the section on split sampling, analysis locations in the same cluster might be spatially correlated. Testing a case against a model derived from the other cases in the same cluster may not yield an entirely independent assessment. The modified jackknife technique requires discarding all cases in one of the k clusters, building a model with the cases in the remaining k-1 clusters, and testing the model on the data in the discarded cluster. This procedure is repeated, with data in each cluster in turn being reserved as the test cases, until k models have been developed and data in each of the clusters have been tested.

When this jackknife method was applied to the Glade Park model data, 36 models, each constructed by eliminating locations in a different sampling unit cluster, were developed. In Table 8.7 the original model, $L(0)$, based on all 314 of the sites and nonsites is given first, followed by the 36 models derived by leaving out the locations in a single cluster. The performance rates found by applying the kth model

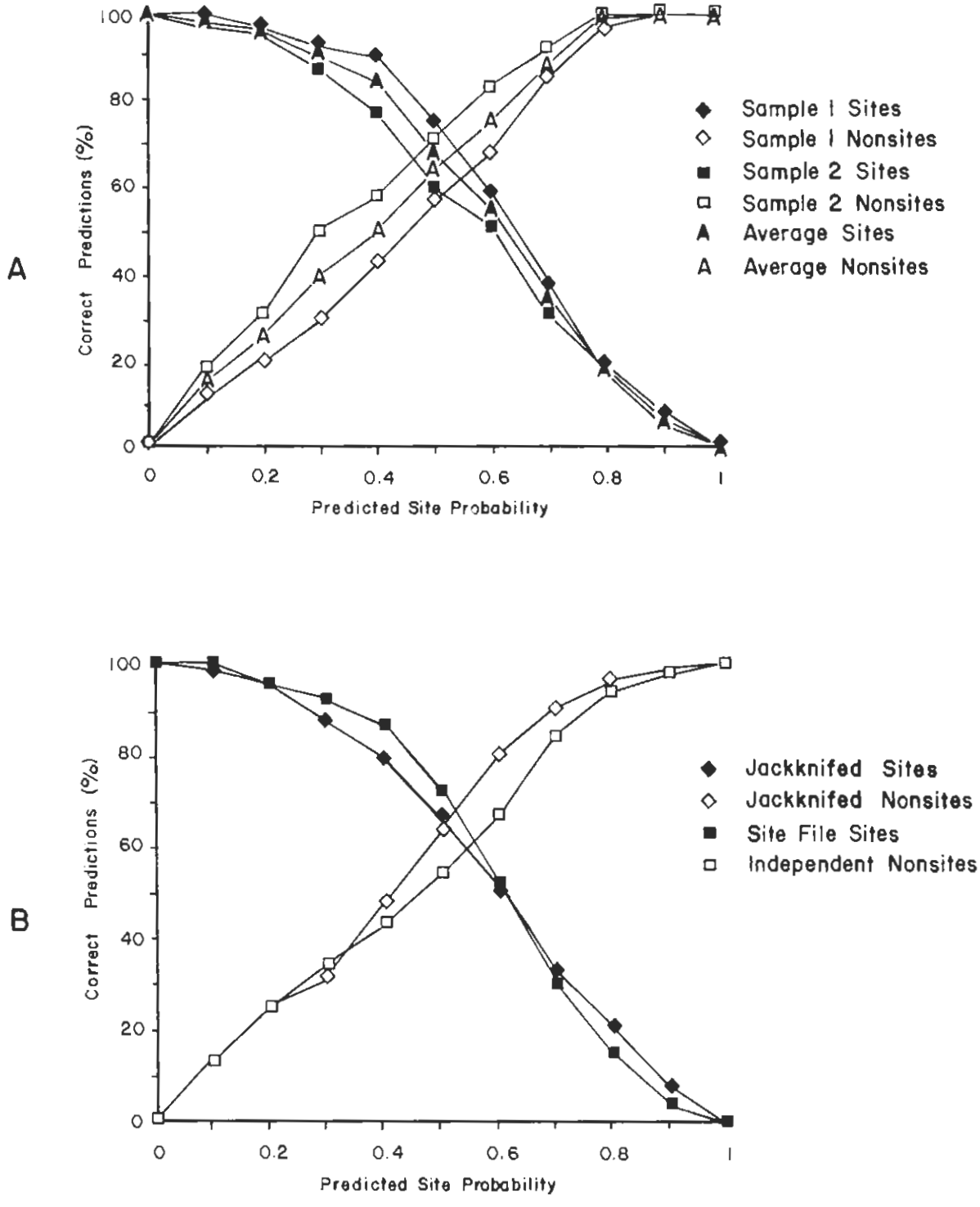


Figure 8.12. Glade Park model performance curves: (A) split-sampled models, (B) jackknifed performance and jackknifed model applied to independent data.

TABLE 8.7.

Original nine-variable Glade Park site-location model (L[0]), 36 jackknifed models, and final composite jackknifed model (L[*])

Model	X_1	X_2	X_7	X_9	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}	*Intercept*
L(0)	-0.002240	-0.013300	-0.000602	-0.040100	0.005180	0.007380	-0.005820	-0.018300	0.000804	0.158000
L(1)	-0.002210	-0.013400	-0.000339	-0.040500	0.005250	0.007270	-0.005760	-0.018500	0.000812	0.146000
L(2)	-0.002120	-0.013200	-0.001070	-0.039900	0.005440	0.007980	-0.005790	-0.017800	0.000821	0.072000
L(3)	-0.002310	-0.013900	-0.000673	-0.040900	0.005060	0.007440	-0.005360	-0.018300	0.000769	0.185000
L(4)	-0.002280	-0.016400	-0.000876	-0.041900	0.006020	0.006960	-0.006450	-0.017600	0.000529	0.423000
L(5)	-0.002070	-0.013000	-0.000581	-0.041000	0.004900	0.006280	-0.006380	-0.018800	0.000816	0.424000
L(6)	-0.002250	-0.012000	-0.000578	-0.039900	0.005340	0.007600	-0.005960	-0.017800	0.000838	0.062000
L(7)	-0.002220	-0.013600	-0.001590	-0.042500	0.004070	0.007210	-0.006870	-0.018400	0.000796	0.550000
L(8)	-0.002210	-0.011800	-0.000620	-0.045100	0.006220	0.006870	-0.006790	-0.020300	0.001120	0.304000
L(9)	-0.002440	-0.012900	-0.000511	-0.038900	0.005590	0.007700	-0.005890	-0.018300	0.000798	0.048000
L(10)	-0.003120	-0.013000	-0.000548	-0.039900	0.005340	0.007150	-0.006400	-0.018600	0.000812	0.337000
L(11)	-0.001950	-0.012700	-0.000607	-0.038300	0.004390	0.007670	-0.005780	-0.019200	0.000883	0.079000
L(12)	-0.001900	-0.014400	-0.000612	-0.038400	0.005720	0.007570	-0.006790	-0.018100	0.000737	0.158000
L(13)	-0.002350	-0.012900	-0.000559	-0.042100	0.006530	0.006940	-0.005830	-0.017600	0.000723	0.174000
L(14)	-0.002290	-0.012600	-0.000629	-0.039300	0.003800	0.007470	-0.005750	-0.020400	0.000977	0.201000
L(15)	-0.002310	-0.014400	-0.000546	-0.042800	0.005290	0.007010	-0.005800	-0.016200	0.000402	0.283000
L(16)	-0.002200	-0.013100	-0.000579	-0.038600	0.005000	0.007030	-0.005380	-0.019000	0.000925	0.183000
L(17)	-0.001940	-0.014000	-0.000731	-0.035800	0.006350	0.007500	-0.004890	-0.019400	0.000962	-0.130000
L(18)	-0.001770	-0.010400	-0.000272	-0.042100	0.004440	0.007910	-0.004890	-0.019500	0.001040	-0.145000
L(19)	-0.001990	-0.011000	-0.000637	-0.036600	0.001250	0.007510	-0.005370	-0.021400	0.001120	0.233000
L(20)	-0.002140	-0.012600	-0.000615	-0.039500	0.005180	0.007250	-0.006220	-0.018700	0.000830	0.138000
L(21)	-0.002230	-0.013400	-0.000689	-0.040300	0.004640	0.008040	-0.006320	-0.016700	0.000863	0.126000
L(22)	-0.002260	-0.015800	-0.000606	-0.040700	0.006180	0.007680	-0.005720	-0.018200	0.000657	0.102000
L(23)	-0.002450	-0.014000	-0.000682	-0.040800	0.005220	0.007770	-0.004560	-0.018900	0.000819	0.088000
L(24)	-0.002270	-0.013700	-0.000464	-0.043700	0.006010	0.007650	-0.005770	-0.019200	0.000828	0.105000
L(25)	-0.002270	-0.012200	-0.000549	-0.039600	0.005250	0.007390	-0.005330	-0.017400	0.000736	0.052000
L(26)	-0.002280	-0.013000	-0.000580	-0.041700	0.005400	0.007620	-0.006000	-0.017000	0.000758	0.158000
L(27)	-0.004670	-0.013100	-0.000596	-0.039100	0.004880	0.007280	-0.005540	-0.018300	0.000799	0.084000
L(28)	-0.002360	-0.013400	-0.000682	-0.041100	0.004890	0.006940	-0.005290	-0.018000	0.000783	0.266000
L(29)	-0.002260	-0.013100	-0.000501	-0.038900	0.006380	0.007470	-0.005630	-0.018200	0.000848	-0.004000
L(30)	-0.002300	-0.014200	-0.000493	-0.040500	0.006760	0.007490	-0.005910	-0.017700	0.000703	0.042000
L(31)	-0.003060	-0.016800	-0.000709	-0.038600	0.004960	0.007440	-0.005850	-0.017300	0.000508	0.331000
L(32)	-0.002770	-0.012100	-0.000669	-0.039900	0.004610	0.007340	-0.006010	-0.017800	0.000786	0.188000
L(33)	-0.002240	-0.013000	-0.000621	-0.039500	0.004800	0.007340	-0.005530	-0.018100	0.000791	0.153000
L(34)	-0.002320	-0.013500	-0.000648	-0.039200	0.004840	0.006850	-0.005640	-0.017700	0.000734	0.276000
L(35)	-0.002100	-0.012700	-0.000528	-0.039900	0.005040	0.007100	-0.006010	-0.018400	0.000822	0.162000
L(36)	-0.002090	-0.012700	-0.000587	-0.040200	0.005190	0.008000	-0.006270	-0.018100	0.000846	0.038000
L(*)	-0.001700	-0.014100	-0.000014	-0.036100	0.005420	0.007340	-0.005620	-0.016500	0.000758	-0.040000

to the discarded cases of the kth cluster are given in Figure 8.12b. At the $p = 0.5$ cutoff point (Figure 8.12b) about 66 percent of the sites and 64 percent of the nonsites are correctly classified, compared with 70 and 66 percent, respectively, for the initial model (Figure 8.11b and Table 8.6). At the $p = 0.3$ cutoff the jackknifed data (Figure 8.12b) suggest that 87 percent of the sites and 32 percent of the nonsites are correctly classified, in contrast to 91 and 38 percent, respectively, for the initial model (Figure 8.12b). Thus, the jackknife suggests moderate decreases in model performance, and these rates may be taken as better estimates of "true" model performance rates. Similar results have been noted elsewhere (e.g., Campbell 1981).

The jackknifed site location model, created by taking a weighted average of the coefficients of the 36 individual models, is given as $L(*)$ in the last line of Table 8.7 (see Mosteller and Tukey 1977:152).

Completely Independent Samples

It was indicated above that one of the most reliable ways to test the performance of a site location model is to apply it to data from a second, independent survey of random sampled data. Such data were not available in Glade Park, but the existing site files, which contain information on many hundreds of known sites, provide a large body of independent site location information. Site forms on file at the local BLM office were carefully screened for quality of information, particularly with regard to accutate locational data (see Chapter 7). A simple random sample of 50 sites that represented a well-spread distribution of sites from throughout the Glade Park region was selected (see Kvamme 1983c for details). A control group of nonsite locations was also chosen so that model performance could be assessed. These nonsites were not necessarily selected from previously surveyed regions and thus actually represent the "environment at large" (rather than true nonsites), but they may still be referred to as nonsites. As discussed earlier, when the prior or chance probability of a site is very low in a region (and in this case the site prior probability has been estimated to be as low as $P[S] = 0.02$; see below), nonsites can be selected at random from throughout a study region regardless of whether or not the locations have been field inspected. The advantage of this procedure is that better estimates of nonsite variation can be obtained than if the nonsites were restricted to a limited number of surveyed clusters. The disadvantage is that some small percentage (here about 2 percent) of the nonsites are misclassified because they are really sites. In the present case 87 nonsites were selected at points located at the center of each of 87 randomly selected sections throughout the region (on a chance basis, only one or two of them should fall on sites).

The jackknifed site location model (last line of Table 8.7) was applied to measurements performed at the 50 independent site and 87 independent nonsite locations. The results of this test, superimposed on the jackknifed results (Figure 8.12b), are very supportive of the performance rates determined by other means.

Statistical Tests

Model classification performance indications are often unreliable owing to a failure to meet the various assumptions of the model used and particularly when the same data are used to build a model and to assess it. In making a statistical assessment of a model's performance, it is much safer to use independent test samples. In other spatially oriented disciplines, statistical significance of model predictions and confidence limits around predictions are commonly determined through the use of independent test samples (e.g., Hay 1979; Rosenfield et al. 1982; Schowengerdt 1983:109–195).

The most common performance assessment of a classification model involves determination of accuracy rates (percent correct statistics). The following sections present a significance test for model classification results and procedures for establishing confidence limits around percent correct statistics obtained when a model is applied to independent test samples. Also presented is a graphic technique for assessing the goodness of fit of a model to the empirical data, which offers an alternative to accuracy rate statistics. Associated with this technique is a significance test that *is* appropriate for application to the same data set from which the locational model was derived. Finally, a sequential analysis approach is presented that minimizes the size of independent test samples needed to test a model by requiring the collection of new data only until a decision about model performance can be reached.

Testing the Significance of Model Classification Results

When an archaeological locational model is applied to independent test samples in a two-class problem (e.g., samples of sites and nonsites), the resulting classification can be statistically assessed through a relatively simple chi-square test for differences in classification probabilities. This test assumes that *(a)* independent test samples from both classes (populations) are being used, *(b)* the test samples are random samples, *(c)* the two samples are mutually independent (i.e., the locations in the site sample really have sites and the locations in the nonsite sample do not have sites), and *(d)* the locations can be unambiguously assigned by a model (decision rule) to either of the classes. The data are arranged in a 2 by 2 contingency table, as shown in Table 8.8a.

A one-tailed test is most appropriate since we are testing for direction in the table, i.e., we are testing whether the model has some utility for making correct classifications. The null hypothesis states that the probability that a location belonging to the site-present population will be classified by the model to that population is less than or equal to the probability that a location from the site-absent population will be classified to the site-present class. Rejection of the null hypothesis implies acceptance of the alternative—that a location from the site-present population has a greater probability of being classified to that population than does a location from the site-absent population, indicating that the model has some predictive utility. The test statistic

$$T = \frac{N\,(ad-bc)^2}{n_1\,n_2\,(a+c)\,(b+d)}$$

(see Table 8.8a for explanation of symbols) is compared against the $(1-2\alpha)$ quantile of the chi-square distribution with one degree of freedom. If T exceeds that value, the null hypothesis may be rejected (see Conover 1971:141–146).

TABLE 8.8.

Assessment of model classification results: (a) set-up for a 2 by 2 table; (b) classification results of jackknifed model applied to independent Glade Park data (at $p = 0.4$ cutoff point); (c) goodness-of-fit test data with fixed cutoff points applied to data used to establish initial Glade Park model

A. Table Set-Up

	Model Class Assignment		
	Class 1 (site present)	*Class 2 (site absent)*	
True Class Membership			
Population 1 (site present)	a	b	n_1
Population 2 (site absent)	c	d	n_2
			$N = n_1 + n_2$

B. Classification Results

	Model Class Assignment		
	Class 1 (site present)	*Class 2 (site absent)*	*Total*
True Class Membership			
Actual site present	43 (86%)	7	50
Site absent (background)	50	37 (43%)	87
			137

C. Goodness of Fit

	Site		*Nonsite*		
Model Cutoff Points	*Expected*	*Observed*	*Expected*	*Observed*	*Total*
0 – 0.0833	0.6144	0	17.3856	18	18
0.0833 – 0.1667	2.3232	4	16.6768	15	19
0.1667 – 0.25	3.9014	5	15.0986	14	19
0.25 – 0.3333	8.5350	8	20.4650	21	29
0.3333 – 0.4167	11.3486	14	18.6514	16	30
0.4167 – 0.5	16.5794	16	19.4206	20	36
0.5 – 0.5833	19.0815	16	15.9185	19	35
0.5833 – 0.6667	24.4641	23	14.5359	16	39
0.6667 – 0.75	25.3594	24	10.6406	12	36
0.75 – 0.8333	19.5763	19	5.4237	6	25
0.8333 – 0.9167	16.6168	19	2.3832	0	19
0.9167 – 1.0	8.5999	9	0.4001	0	9
Total		157		157	314

The independent Glade Park test data results from the previous section can be used to illustrate application of this test. The independent sample of site locations was taken from existing site file information at a local BLM office (this sample was discussed above); because it is possible that survey biases might be reflected in this sample, in practice it would be more desirable to use sites obtained from an independent field survey conducted under a random sampling design. The independent sample of nonsite locations is actually a sample of locations taken at random from the background environment at large (this sample was also described above). Without field checking, there is no way of knowing for certain whether or not a particular location in this sample contains a site; an estimate of the base rate or a priori chance of a site occurring at a location in the region (see below), however, indicates that approximately 94–98 percent of this sample should not contain sites. Although the third assumption listed above technically is violated, the performance of the test should be modified only slightly given the low rate of site occurrence (the principal effect will be to make acceptance of the null hypothesis more likely through a reduction in the apparent significance of the model). Note that even if a sample of actual nonsite locations were obtained, there would always be some uncertainty about the absence of sites from all sample locations owing to the possibility of sites having been missed during survey and to the potential presence of buried sites.

The independent test data indicate that at the $p = 0.4$ cutoff point approximately 86 percent of the locations with sites and 43 percent of the locations without sites are correctly classified by the Glade Park jackknifed model (Figure 8.12b), which produces the 2 by 2 structure shown in Table 8.8b. When computed using these data, the test statistic yields

$$T = \frac{137\,[(43)(37) - (7)(50)]^2}{(50)(87)(43 + 50)(7 + 37)} = 11.853$$

At a level of significance of 0.001 the null hypothesis will be rejected if T exceeds 9.549 (from a table of the chi-square distribution with one degree of freedom). It is therefore rejected in the current case, which suggests that the model has some predictive utility at the $p = 0.4$ cutoff point. (A common complaint with contingency table tests in archaeology is that a significant result might be due to only one cell with a large deviation from expectation. In this testing context, however, if most of the test statistic is due to one cell it means that either the model does a better job than chance at classifying sites or that the model does better than chance at classifying nonsites. In either case we win because the model offers some gain over pure chance.)

Before leaving the subject of testing model accuracy rates, it should be noted that a number of additional procedures currently being examined in other disciplines warrant investigation by archaeologists. These include specialized analysis of variance techniques (Rosenfield 1981) and the set of methods known as discrete

multivariate analysis (Congalton et al. 1983). Both of these approaches offer the potential for significance testing of individual and overall classification results in tables larger than 2 by 2, making them suitable for multiclass modeling problems (e.g., models for multiple site-type classes).

Establishing Confidence Limits Around Accuracy Probabilities

Since the classification of a test location by a model or decision rule is either right or wrong (i.e., a site or nonsite location is correctly identified or it is not), the correctness of a classification assignment at each location represents a binomial population. The Glade Park independent data test results (Table 8.8b) indicate that 86 percent of the site locations and 43 percent of the nonsite locations should be correctly classified by the jackknifed model (at the 0.40 cutoff). These percent correct statistics, which represent estimated mean probabilities of correct classification (when divided by 100), can be considered random variables with a binomial probability distribution. Associated levels of statistical error can be found in tables or graphs of confidence limits of the mean of a binomial distribution (e.g., Conover 1971:380–381; Hord and Brooner 1976). Hord and Brooner (1976:672) give the following as the approximate $100(1 - \alpha)$ percent confidence interval for p, the proportion of successes, given n trials.

$$\frac{p + \left(\frac{z^2_{\alpha/2}}{2n}\right) \pm z_{\alpha/2}\sqrt{\frac{(p)(1-p) + (z^2_{\alpha/2}/4n)}{n}}}{1 + (z^2_{\alpha/2}/n)}$$

When the Glade Park results for the site class are used, the proportion of sites correctly classified by the model is $p = 0.86$ and $n = 50$. For a 95 percent confidence interval, a table of the normal distribution (found in any statistical text) gives $z_{\alpha/2} = z_{.025} = 1.96$. The limits of the 95 percent confidence interval become

$$p_{lower} = \frac{0.86 + \left(\frac{1.96^2}{2(50)}\right) - 1.96\sqrt{\frac{(0.86)(1 - 0.86) + 1.96^2/(4[50])}{50}}}{1 + (1.96^2/50)} = 0.74$$

$$p_{upper} = \frac{0.86 + \left(\frac{1.96^2}{2(50)}\right) + 1.96\sqrt{\frac{(0.86)(1 - 0.86) + 1.96^2/(4[50])}{50}}}{1 + (1.96^2/50)} = 0.93$$

or $P(0.74 \leq p \leq 0.93) = 0.95$. Similar calculations for the nonsite class (with $p = 0.44$, $n = 87$) yield a 95 percent confidence interval of $0.34 \leq p \leq 0.54$.

In theory, such confidence intervals indicated that, 95 percent of the time, independent test samples should yield proportion correct statistics (p) between these limits. In other words, if we had numerous independent test samples of known site-present locations (and $n = 50$), in about 95 percent of those samples the proportion of sites correctly classified by the model would be between 0.74 and 0.93. The range produced by these limits thus gives a more realistic idea of true model performance.

The width of a confidence interval at a given level of significance is a direct function of the size of the sample used to compute the interval. Hence, it is important to obtain large test samples in order to produce narrower confidence limits. To illustrate, if we increase n to 150 for the above site-class 95 percent confidence interval (leaving $p = 0.86$), we obtain $0.80 \leq p \leq 0.91$. Increasing n to 300 gives $0.82 \leq p \leq 0.89$. Upper and lower confidence limit values can be inserted into other formulas (e.g., the gain statistic or those shown in the base rate probabilities section, below) to assess upper and lower bounds on other dimensions of model performance. Confidence intervals are not restricted to 2 by 2 tables but may be applied to results obtained from tables of any size (e.g., in problems with multiple site types). Parker (1985) illustrates use of the Poisson distribution when estimated mean archaeological probabilities are extremely low (e.g., $p \leq 0.05$).

Assessing Model Goodness of Fit

Parker (1985) presents an alternative approach for assessing archaeological model performance that does not focus on percent correct statistics but compares observed with predicted probabilities of site presence. In this approach, which yields a graphical result, a probability scale (i.e., a scale ranging from 0 to 1) of site presence is divided arbitrarily into multiple groups or intervals (e.g., $0 \leq p \leq 0.02$; $0.02 \leq p \leq 0.06$; $0.06 \leq p \leq 0.10$, etc.; Parker 1985:192). Using predicted site probability values estimated at sample locations by a logistic regression model, the number of known sites and the number of known nonsites that fall in each interval is determined, and the proportion of the total number of locations that are sites is calculated. This proportion is taken as an estimate of the *observed* probability of site presence in each interval. *Expected* probabilities for each interval are calculated simply as the group midpoint value (e.g., the midpoint of the interval $0.02 \leq p \leq 0.06$ is 0.04). The observed and expected pairs for each interval are then plotted on a graph that can be used to assess model goodness of fit. If the plotted points follow a line with an intercept of 0 and a slope of 1 (a 45° angle), the model offers a good fit (Parker 1985:190–192).

A problem with Parker's method is that it is largely subjective; goodness of fit must be determined through a visual assessment of how well the observed and expected values follow a straight line. There is no associated significance test. Moreover, the specific group intervals used in Parker's application were of varying width and were apparently formed during analysis to maximize agreement between observed and expected values. This tactic may have been necessary, however, owing to the extremely small sample size (30 sites) under investigation.

Medical researchers have developed a remarkably similar approach for assessing the goodness of fit of predictive models based on logistic regression. In this approach, though, the grouped probability intervals are specified *prior* to the analysis, allowing more objective results, and an associated significance test is available. Of primary importance for archaeological purposes is that the test may appropriately be applied to the data set from which a model was derived, forming an important tool for screening out useless models prior to further testing. This approach to goodness-of-fit assessment also utilizes the probability of site presence estimated for a location by a statistical model, p. Intervals of equal width are formed, e.g., 0–0.1, 0.1–0.2, . . ., 0.9–1.0, and locations (cases) are assigned to the intervals on the basis of p. If the model has predictive utility, then the p for locations with sites should fall into the upper intervals. The observed number of locations with sites (o_s) is compared with an expected number of locations with sites (e_s) for each interval. The latter is usually calculated as the sum of the estimated p-values for all locations in a particular interval. More explicitly,

$$o_{sk} = \sum_{i \in k} y_i$$

$$e_{sk} = \sum_{i \in k} p_i$$

where $k = 1, \ldots, g$ intervals; o_{sk} is the observed number of sites in the kth interval, e_{sk} is the expected number of sites in the kth interval, $i \in k$ denotes that the ith case is a member of the kth interval, and y_i is coded 1 for sites and 0 for nonsites. In other words, for a particular k^{th} interval (e.g., 0.8–0.9), o_{sk} represents the observed count of sites having site-class p-values that fall in that interval; e_{sk} is simply the sum of the site-class p-values for all locations, site and nonsite, that fall in that interval (Lemeshow and Hosmer 1982). As in Parker's (1985) application, the g pairs of observed and expected values may be plotted, allowing a subjective assessment of goodness of fit when compared with a line with an intercept of 0 and a slope of 1 (Brand et al. 1976).

The comparison of observed and expected site frequencies has been developed into a statistical test for goodness of fit by Lemeshow and Hosmer (1982). Since considerable information is lost when only the site group is considered, a more powerful testing procedure is made possible by considering observed and expected frequencies for site and nonsite classes simultaneously. Observed and expected frequencies for the nonsite group are calculated as follows:

$$o_{nk} = \sum_{i \in k} (1 - y_i)$$

$$e_{nk} = \sum_{i \in k} (1 - p_i)$$

where o_{nk} is the observed number of nonsites in the kth interval and e_{nk} is the expected number of nonsites in the kth interval. The statistic developed by Lemeshow and Hosmer (1982:97) is

$$H_g^* = \sum_{k=1}^{g} \left[\frac{(o_{sk} - e_{sk})^2}{e_{sk}} + \frac{(o_{nk} - e_{nk})^2}{e_{nk}} \right]$$

where the summation is from $k = 1, \ldots, g$ intervals.

The site location model used to illustrate this test for goodness of fit is the initial nine-variable Glade Park logistic regression model given in an earlier section of this chapter. Results of applying this model to the same data that were used to construct the model are tabulated in Table 8.2. A requirement of the test is that the number of intervals (g) should be greater than j+1, where j is the number of predictor variables used by the model (Lemeshow and Hosmer 1982:96). In the present case, $j + 1 = 10$; hence, 12 intervals, with constant widths of 0.0833, are used. The observed and expected frequencies tabulated for the site and nonsite groups in Table 8.8c are used to calculate

$$\begin{aligned} H_g^* = {} & (0-0.61)^2/0.61 + (4-2.32)^2/2.32 + (5-3.90)^2/3.90 + (8-8.54)^2/8.54 \\ & + (14-11.35)^2/11.35 + (16-16.58)^2/16.58 + (16-19.08)^2/19.08 \\ & + (23-24.46)^2/24.46 + (24-25.36)^2/25.36 + (19-19.58)^2/19.58 \\ & + (19-16.62)^2/16.62 + (9-8.60)^2/8.60 + (18-17.39)^2/17.39 \\ & + (15-16.68)^2/16.68 + (14-15.10)^2/15.10 + (21-20.47)^2/20.47 \\ & + (16-18.65)^2/18.65 + (20-19.42)^2/19.42 + (19-15.92)^2/15.92 \\ & + (16-14.54)^2/14.54 + (12-10.64)^2/10.64 + (6-5.42)^2/5.42 \\ & + (0-2.38)^2/2.38 + (0-0.40)^2/0.40 \\ = {} & 8.28 \end{aligned}$$

The distribution of this statistic is approximated by a chi-square distribution with g-2 = 10 degrees of freedom. At a level of significance of $\alpha = 0.05$ the null hypothesis of a good fit can be rejected if H_g^* exceeds 18.31. Since H_g^* is smaller than that value, we can accept the null hypothesis. In fact, the null hypothesis could be accepted at $\alpha = 0.5$.

A similar goodness-of-fit test is presented by Costanzo et al. (1982). This test focuses on residuals rather than predicted probabilities.

Sequential Methods

An approach to model testing that potentially requires smaller test samples was presented in an archaeological study by Limp and Lafferty (1981:226–229). The approach utilizes a sequential probability ratio test or SPRT (Dixon and Massey 1957:304–310; Wetherill 1975). The SPRT requires the collection of new sample data, but only until a decision about a model's performance can be reached. That is, the sequential method does not require the collection of more observations than are necessary to make a decision. This approach can be beneficial for model testing since it offers the potential for reduced amounts of additional survey and, therefore, lower costs.

The SPRT allows a decision between two simple hypotheses. Suppose there is interest in the parameter Θ, the true site density in a low site probability stratum established by an archaeological model. We wish to test the null hypothesis that the true site density equals some specified level, $\Theta = \Theta_0$, against the alternative hypothesis that the true site density equals some other specified level, $\Theta = \Theta_1$. The SPRT decides in favor of either Θ_0 or Θ_1 on the basis of sample observations. If Θ_0 is true, we would like to decide in its favor with a probability of $1-\alpha$ or greater; if Θ_1 is true, we would like to decide for Θ_1 with a probability of $1-\beta$ or greater.

To illustrate, a predictive archaeological model developed in southern Arkansas yielded a low site probability stratum that was mapped throughout the entire region of study (Limp and Lafferty 1981). The unit of analysis was a 4 ha grid unit (i.e., a land parcel 200 m on a side); the entire region was gridded into more than 3000 such units. Based on the sample data used to establish the model it was estimated that in the low probability stratum the proportion of all grid units with sites was only 0.009. Limp and Lafferty (1981:227) were willing to accept the model if the true proportion (Θ) of units with sites in the low probability stratum really was 0.009 or less. They therefore established an SPRT to test with independent data the null hypothesis that the true site proportion is $\Theta = 0.009$ (setting the probability of falsely rejecting the null hypothesis at $\alpha = 0.10$). Their alternative hypothesis was that the true portion of units with sites was $\Theta_1 = 0.025$, an arbitrary proportion that they deemed would yield an unacceptably high number of site-present grid units in the low probability stratum. (They set the probability of falsely accepting the null hypothesis, i.e., accepting Θ_0 when Θ_1 is really true, at $\beta = 0.10$.) Thus, their sequential test was established in order to decide whether to accept $\Theta_0 = 0.009$ (or less) or an alternative, $\Theta_1 = 0.025$ (or greater), as the true site proportion.

The SPRT requires that observations (grid units) be made, by random selection, one at a time. After each observation, one of three decisions is made: *(a)* accept the null hypothesis ($\Theta_0 = 0.009$), *(b)* reject the null hypothesis by accepting the alternate hypothesis ($\Theta_1 = 0.025$), or *(c)* make an additional observation. The test offers an easy-to-use graphic counterpart established by the following formulas. An upper limit is given by

$$(G_s)ln(\Theta_1/\Theta_0) + (G_n)ln[(1-\Theta_1)/(1-\Theta_0)] = ln[(1-\beta)/\alpha]$$

and a lower limit by

$$(G_s)ln(\Theta_1/\Theta_0) + (G_n)ln[(1-\Theta_1)/(1-\Theta_0)] = ln[\beta/(1-\alpha)]$$

where G_s is the number of grid units currently inspected with sites and G_n is the number of grid units with no sites. Inserting values defined above yields

$$(G_s)ln(0.025/0.009) + (G_n)ln[(1-0.025)/(1-0.009)] = ln[(1-0.10)/0.10]$$

and

$$(G_s)ln(0.025/0.009) + (G_n)ln[(1-0.025)/(1-0.009)] = ln[0.10/(1-0.10)]$$

yielding, after simplification, the following respective upper and lower limit equations:

$$1.0217(G_s) - 0.0163(G_n) = 2.1972$$

and

$$1.0217(G_s) - 0.0163(G_n) = -2.1972$$

These limits can be plotted as parallel lines in a graph by finding two points for each and drawing a line through them. Setting $G_n = 0$ gives 2.15 for the upper limit and -2.15 for the lower limit. Setting $G_n = 200$ gives 5.34 for the upper limit and 1.06 for the lower limit. The upper and lower limit lines plotted through these points are graphed in Figure 8.13.

In a model testing context, graphs such as Figure 8.13 are established prior to testing. During the survey, the result for each test observation (grid unit) is plotted by drawing a line one unit to the right if the observation does not contain a site and one unit upward if the observation contains a site. Sampling is continued until the plotted line crosses the upper or lower limit, at which time a decision is reached concerning the acceptance or rejection of the model. If the true proportion of units with sites is exactly equal to Θ_0, then the null hypothesis will be accepted approximately $100(1-\alpha)$ percent of the time (upon repeated testing trials); if the true proportion of units with sites is exactly equal to Θ_1, then the null hypothesis will be accepted about $100(\beta)$ percent of the time; if the true proportion of units with sites is between Θ_0 and Θ_1, then the null hypothesis will be accepted between $100(1-\alpha)$ and $100(\beta)$ percent of the time, the percentage of acceptance decreasing progressively from $100(1-\alpha)$ to $100(\beta)$ as the true proportion increases from Θ_0 to Θ_1.

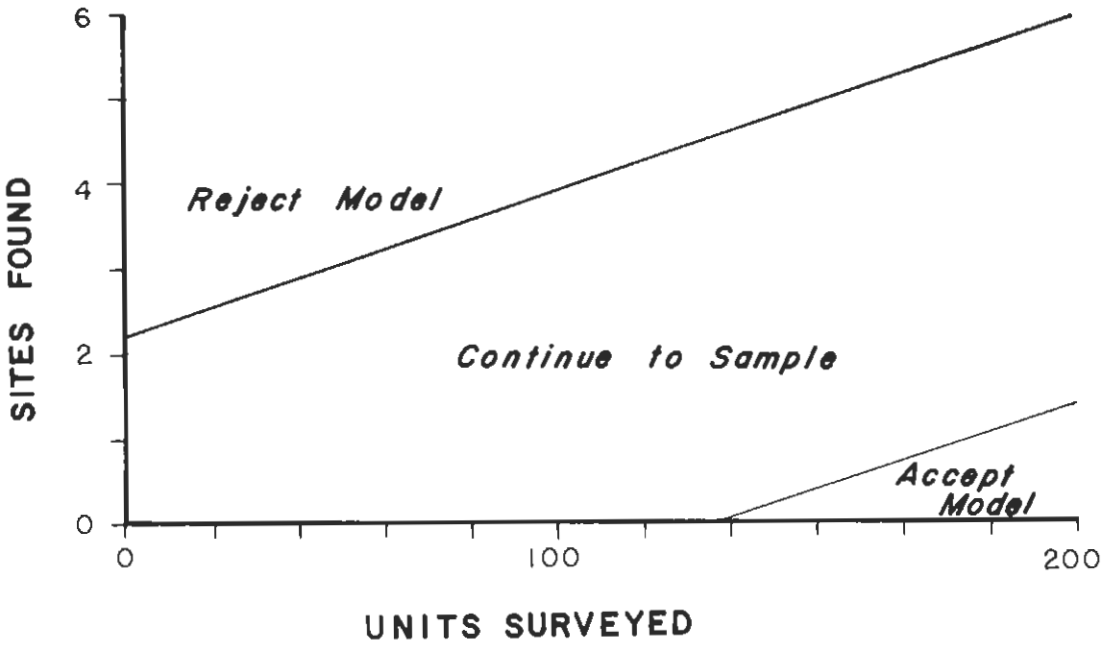


Figure 8.13. Sequential sampling design for a southern Arkansas study. For a given number of units surveyed, if the number of sites encountered exceeds the upper limit, the site density expected by a predictive model is exceeded and the model may be rejected (after Limp and Lafferty 1981:227).

In Limp and Lafferty's (1981) application (Figure 8.13), it is readily apparent that if the first 135 grid units sampled did not contain sites, the model could be accepted immediately (i.e., $\Theta = \Theta_0 = 0.009$). On the other hand, discovery of only three units with sites for sample sizes under 53 would be cause for immediate rejection of the model and acceptance of $\Theta = \Theta_1 = 0.025$. The space between the acceptance and rejection regions represents an "inconclusive" range where neither decision can be reached.

Formulas for estimating the average sample size needed to arrive at a decision are given by Dixon and Massey (1957:309–310). Application of these formulas to the Limp and Lafferty data yields $n = 253$ when the true proportion of sites is $\Theta_0 = 0.009$; $n = 183$ when the true proportion is $\Theta_1 = 0.025$; and $n = 290$ when the true proportion is between Θ_0 and Θ_1 (this latter figure is approximately the maximum average sample size).

It should be emphasized that the Limp and Lafferty (1981) example illustrates a rather extreme application of sequential methods because they focused on such low probabilities (i.e., $\Theta_0 = 0.009$). In any statistical procedure dealing with rates and proportions, a number of problems arise when estimated probabilities are very high or very low. First, since the estimated probabilities are based on relative frequencies derived from empirical data, very large samples are needed for reliable estimates when the relative frequencies are extreme (e.g., less than 0.05 and greater than 0.95). Limp and Lafferty (1981) derived $\Theta_0 = 0.009$ by finding two sites in only 235 units in their initial sample. A change of only two sites in either direction would have caused Θ_0 to range between 0 and 0.017, substantially altering the structure of the sequential test given above or even preventing its use (in the $\Theta_0 = 0$ case). A sample size of several thousand would be needed for a reliable estimate of $\Theta_0 = 0.009$. Second, extreme estimated probabilities in a sequential test require that large samples be examined before a decision regarding acceptance or rejection of Θ_0 can be made. To illustrate, if a more reasonable low probability stratum that contained approximately 20 percent of all sites had been defined, then $\Theta_0 = 0.2$. Suppose that a determination had been made that this stratum could acceptably contain as many as 30 percent of all sites; then $\Theta_1 = 0.3$. The average sample size needed to arrive at a decision (leaving $\alpha = \beta = 0.1$) would be $n = 69$ when the true proportion is $\Theta_0 = 0.2$; $n = 63$ when the true proportion is $\Theta_1 = 0.3$; and $n = 90$ when the true proportion is between Θ_0 and Θ_1 (compare $n = 253$, $n = 183$, and $n = 290$, respectively, for the Limp and Lafferty application above).

Several important assumptions and technical difficulties behind the sequential method limit its practical use. Sequential methods assume complete randomization of sampling units. After each unit is inspected a new decision is made; therefore, the next unit must be chosen at random. This prohibits the typical practice of selecting clusters of units located near one another for each day's work in order to minimize travel. Each unit must be selected at random, and the units must be inspected in random order. This requirement necessarily causes increased effort to be expended in travel to sampling units. This difficulty may be reduced to some extent by selecting sampling units in groups (e.g., groups of 10) rather than individually; this

would allow some flexibility in travel plans. The sequential test would then be assessed after surveys of each group had been completed. The principal effect on the procedure would be to increase the average sample size needed to arrive at a decision by an amount equal to the size of each group (Dixon and Massey 1957:310).

Base Rate Probabilities

Previous sections have presented a number of procedures for assessing the performance of a model through independent samples and significance tests. Before we can fully assess a particular model or understand how well it will work in practice, we must take into account one final domain—the base rate or a priori site and nonsite probabilities, which have been mentioned several times in previous sections. By using these probabilities one can make estimates of the probability of site class membership within a *region* mapped by a model or, alternatively, estimate the probability of site class membership at *specific loci* within a region of study.

Archaeological sites are rare phenomena. This can be clearly demonstrated by examining the a priori probability of site occurrence within a region—the purely chance probability of site presence considering no other information. This probability is usually extremely low, ranging in the vicinity of 1 to 5 percent or even much less. This probability can be estimated as

$$P(\text{site}) = P(S) = \frac{\text{total area covered by known sites}}{\text{total area surveyed}}$$

The total area covered by known sites is most accurately estimated by measuring site area in the field or by determining the area of the dots and polygons usually used to record site locations on maps. If a small grid (e.g., one of 50 by 50 m cells) is superimposed over the study region and the number of grid cells that contain cultural remains are counted, then $P(S)$ can be estimated simply by dividing the total number of cells with sites by the total number of field-inspected cells. Reliable estimation of $P(S)$ always requires fairly large samples. It is important to note that the gridding method can cause an overestimate of $P(S)$ when a large grid size is used. A large cell is more likely to contain a site than a smaller one, and this causes the relative number of cells with sites to increase while the total number of cells is decreased.

The Glade Park data can be used once again to reveal that 157 of the 2432 surveyed analysis units (each measuring 1 ha) contain sites, yielding an estimate of $P(S) = 157/2432 = 0.065$. Most of the sites discovered, however, were very small lithic scatters covering an area much smaller than a hectare, which suggests that the above figure is an overestimate. Examination of the site records indicates that the 157 sites occupy a total area estimated at about 538,000 m², or an average size of less than 3500 m² (compared to 10,000 m² in a hectare). Since 38 quarter-sections occupy approximately 25 million square meters, a better estimate of the actual base rate

probability of archaeological site presence in Glade Park might be $P(S)$ = 538,000/25,000,000 = 0.021. Since Glade Park is one of the archaeologically richest areas in Colorado, this figure is relatively high.

Incorporation of prior probabilities into a classification model will decrease the *overall* rate of misclassification (Morrison 1976:235), but when the prior probability of one group is extremely low (as in the archaeological case) the error rate for the low-probability class is increased substantially by this procedure (Morrison 1969:160; Overall and Klett 1972:263). The extreme magnitude of the prior probabilities in such cases "overpowers" the estimated probabilities that are conditional on environmental and other data, with the effect that the final model essentially utilizes only the prior information in classifying observations. It is best, therefore, not to include prior probabilities in model development but to reserve them for model performance assessment (see below, however, for a discussion of the use of prior probabilities in estimates of probabilities at specific loci). Some disciplines actually manufacture a priori probabilities, arbitrarily setting $P(S) = 0.9$, for example, in an effort to increase the chance that a rare group of interest will be correctly identified by a predictive classification model (Schowengerdt 1983:43). This procedure is mathematically equivalent to the cutoff point adjustment approach explained in an earlier section.

Estimating Site Probabilities in Regions

Since archaeological sites are a valuable resource, it is more important for archaeological locational models to classify site-present locations correctly than for models to classify site-absent locations correctly. We would like, therefore, to produce models that classify a major proportion of sites correctly, say 90 percent. This can be accomplished using the method of modified cutoff points described above and the nonsite data can be used to indicate the approximate percentages of the study area within which a specified percentage of sites should occur. But in order to determine other dimensions of model performance, such as the site densities that can be expected, we can use prior probabilities with the model performance indications obtained through the cutoff point adjustment approach and Bayes's Theorem (Hays 1981:39–41). More specifically, given an area of a region mapped by a model as site-likely or site-favorable, the following procedures yield an estimate of the probability of site class membership within that modeled region and an estimate of the probability of site class membership outside the modeled region.

To illustrate this procedure the percent correct statistics yielded by applying the jackknifed Glade Park model to the independent test data (Table 8.12b) are used. These data indicate (at a model cutoff point of $p = 0.4$) that approximately 86 percent of the sites should be classified correctly (Table 8.8b). Let S be the event that a site is actually present, and let M be the event that the model indicates that a site is present. We want to find the conditional probability, $P(S|M)$, of site class membership given that the model suggests site presence. If we use the grid-based analysis, the a priori probability of site presence at a location is estimated as $P(S)$ =

0.065 (see above); then $P(S^c) = 0.935$, where S^c indicates the complement of site presence, i.e., site absence. The probability that the model will indicate a site given that a site is actually present is $P(M|S) = 43/50 = 0.86$, and the probability that the model will indicate a site given that a site is *not* present is $P(M|S^c) = 50/87 = 0.575$ (data from Table 8.8b). According to Bayes's Theorem,

$$P(S|M) = \frac{P(M|S)P(S)}{P(M|S)P(S) + P(M|S^c)P(S^c)}$$

$$= \frac{(0.86)(0.065)}{(0.86)(0.065) + (0.575)(0.935)} = 0.094$$

Consequently, in the portion of the study region that this model would map as site-likely (at the $p = 0.4$ cutoff), the probability of site class membership at any location (hectare cell) within the region is $P(S|M) = 0.094$, which is 0.094/0.065 or 1.45 times better than a purely chance model ($P[S] = 0.065$). On the other hand, the probability of site class membership given that the model does not indicate a site is roughly

$$P(S|M^c) = \frac{P(M^c|S)P(S)}{P(M^c|S)P(S) + P(M^c|S^c)P(S^c)}$$

$$= \frac{(0.14)(0.065)}{(0.14)(0.065) + (0.425)(0.935)} = 0.022$$

In the portion of the environment *not* mapped by the model as site-likely the probability of site class membership is only $P(S|M^c) = 0.022$. This suggests that haphazardly throwing darts at a map of the region (a purely chance model) might be three times (0.065/0.022) more probable of indicating a site than the probability produced by the model in this subarea. Moreover, the probability of site class membership in the mapped site-likely region is more than 4.2 times (0.094/0.022) more likely than the probability of a site occurring in the site-unlikely region. (It is emphasized, once again, that these procedures can be extended to problems involving multiple site classes.)

The meaning of these statistics is made clearer by imagining that the Glade Park model (at the $p = 0.4$ cutoff) is mapped over the entire study region (roughly 160,000 ha), much like the mappings in Figure 8.8. About 6.5 percent of these hectare-unit locations ($P[S] = 0.065$), or 10,400 of them, will contain sites, and about 93.5 percent ($P[S^c] = 0.935$) or 149,600 will not (Figure 8.14), as estimated by their base rate chances of occurrence. Of the 10,400 locations that contain sites, the predictive site location model (at the $p = 0.4$ cutoff point) will (as indicated by the independent tests) correctly classify about 86 percent ($P[M|S] = 0.86$) or 8944 as

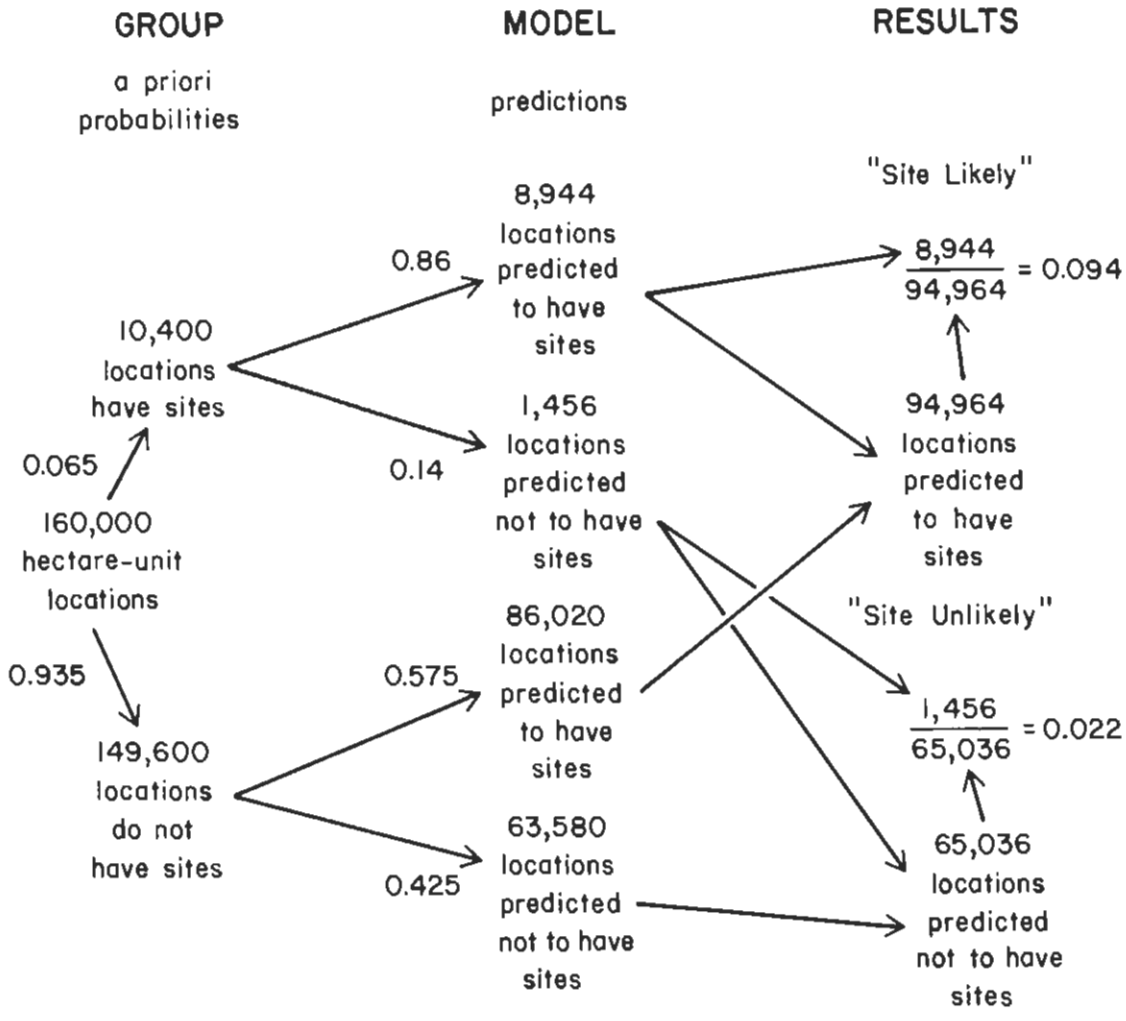


Figure 8.14. Illustration of Bayes's Theorem and the effects of site and nonsite a priori probabilities on Glade Park model performance.

containing sites and will incorrectly classify about 14 percent ($P[M^c|S] = 0.14$) or 1456 as belonging to the site-absent category. Of the 149,600 nonsite locations, 42.5 percent ($P[M^c|S^c] = 0.425$) or 63,580 will be correctly identified and about 57.5 percent ($P[M|S^c] = 0.575$) or 86,020 will be classified as sites. Thus, although the model assigns 8944 + 86,020 = 94,964 locations as site-likely, only 8944 of these actually contain sites, or 8944/94,964 = 0.094 = $P(S|M)$, a roundabout, and hopefully more understandable, presentation of Bayes's Theorem. These calculations are illustrated in Figure 8.14.

It is important to recognize that the predicted 94,964 site-likely hectares of the model can potentially be mapped through use of computer mapping techniques (see above and Chapter 10). About 86 percent of all sites would occur within the approximately 57 percent of the total land area that is mapped by the model as having high site sensitivity. The area *outside* the mapping would form a low-sensitivity zone covering about 43 percent of the land area and would contain only 14 percent of all sites. In fact, 100(63,580)/(63,580+1456) = 97.8 percent of the locations in the low-sensitivity zone would not contain sites. About one location in every 10 would contain a site in the high-sensitivity zone, but only one location in every 45 would contain a site in the low-sensitivity region. These statistics, of course, are based on the model using the $p = 0.4$ cutoff and on accuracy rates obtained from one sample (Table 8.8b). Performance indications such as these will vary depending on the cutoff point and accuracy estimates used.

Estimating Site Probabilities at Specific Loci

Cultural resource managers often wish to estimate the probability of archaeological site class membership given the data measured at a particular location, such as a single hectare grid cell, rather than simply estimating the probability of a site within a larger region, such as a high-sensitivity zone as a whole. Probability estimates for specific loci also require use of the a priori probabilities $P(S)$ and $P(S^c)$. These probabilities are used in conjunction with modifications of the formulas for estimating site probabilities conditional on environmental and other measurements (given in the section "Application Comparison of Quantitative Locational Models" above).

It was demonstrated with empirical test evidence that if the Glade Park jackknifed model (at the $p = 0.4$ cutoff) were to be mapped, the probability of site class membership within the mapped site-likely region would be about 0.095, and the probability of site class membership outside the mapped region would be about 0.022. These estimates, one for the entire area mapped by the model and one for the rest of the study region, respectively, serve as a kind of "average" probability figure for these portions of the study area. In other words, if we know that a location falls somewhere within the region mapped by the model as site-likely, then we can say that the probability of site class membership is about 0.095. This technique makes use only of the knowledge that a location is, or is not, in a modeled region as a whole, mapped at some cutoff point; it does not consider any particular factors, such as environmental characteristics at a particular location.

It is also possible to estimate the probability of site class membership at a specific locus (land parcel) by ignoring the mapping and considering the environmental characteristics of the locus together with the base rate or chance probability of a site (but for complete validity this procedure technically requires that all the assumptions of the classification model used are met). The discriminant analysis described in the application comparison section (above) provides an example. That analysis yielded a discriminant score of $D_i = 2.2009$ for the environmental data measured at the location of site 5LA5364 (Table 8.3). This location's probability of membership in the site class, conditional *only* on the environmental measurements and assuming that the assumptions of the discriminant model were fully met, was estimated as

$$p_i = \frac{e^{-0.5(D_i - D_s)^2}}{e^{-0.5(D_i - D_s)^2} + e^{-0.5(D_i - D_{ns})^2}} = 0.873$$

(recall that $D_s = 0.8304$ and $D_{ns} = -0.1936$). Again, this probability is estimated from the measurements only and does not consider the base rate proportions of sites and nonsites in the area. A modification of this formula to incorporate prior probabilities, $P(S)$ and $P(S^c)$, yields

$$p_i = \frac{P(S)e^{-0.5(D_i - D_s)^2}}{P(S)e^{-0.5(D_i - D_s)^2} + P(S^c)e^{-0.5(D_i - D_{ns})^2}}$$

and the estimated probability of site class membership at this location, incorporating both environmental and base rate data, is approximately $p = 0.323$ (using $P[S] = 0.065$ and $P[S^c] = 0.935$). The lower figure results from the inclusion of the prior information on site proportions and provides a more realistic estimate of anticipated probabilities. Similar modifications of other formulas (e.g., logistic regression) can be found in standard statistical texts.

MODEL REVISION

Analyses described in the previous sections suggested that about 86 percent of Glade Park sites might be predicted correctly by mapping a high site-sensitivity zone that covers approximately 57 percent of the total Glade Park land area. This particular result may not seem very impressive as an illustration of the power of empirical site location models. It was noted earlier, however, that Glade Park contains one of the highest site densities in Colorado. This fact, together with these area performance indications, suggests that Glade Park was a very favorable place for prehistoric peoples to perform activities and, in the process, create archaeological sites. The 57 percent figure suggests that about 57 percent of the land area of

Glade Park contains environmental characteristics that are very similar to characteristics exhibited by known sites (in terms of a partitioning of measurement space, Figure 8.5b). This means that prehistoric peoples had a wide choice of settlement locations within the region. Other regions and studies do not indicate such favorable conditions for prehistoric inhabitants. The Colorado plains study described in earlier sections (also see Kvamme 1984) found that sites were restricted primarily to a narrow zone around major drainages. Statistics obtained through independent testing suggested that about 90 percent of the sites might occur in only 50 percent of the total land area of that study region. In a study in central Utah more than 90 percent of the sites were estimated to occur in about 15 percent of the study region's area (Reed and Chandler 1984:80).

It is through the use of nonsite control data that these area projections can be made. Because many nonsite locations exhibit environmental characteristics identical to those of sites (and thus fall on the site side of the decision boundary in the measurement space), and because they are extremely prevalent, these approximate area calculations can be made. Although much of the (nonsite) environment may possess characteristics similar to those exhibited by known site locations, much of the (nonsite) environment is very dissimilar, which allows the designation of substantial portions of the environment as a low site-sensitivity zone. Thus, at Glade Park 43 percent of the land area could be delineated as having low site sensitivity, a result that would include only about 14 percent of the prehistoric sites within that zone. At present, no method has been demonstrated that can discriminate site-present from site-absent locations in the site-favorable portion of a measurement space. In other words, given that there are many locations in the environment that possess environmental and other characteristics identical to those exhibited at site locations, there presently is no procedure that can differentiate between sites and nonsites with identical environmental and other characteristics.

A projection like "90 percent of the sites will occur in 90 percent of the land area" offers no gain. In assessing whether gain is sufficient, such factors as test sample sizes and confidence interval widths should be considered. If it is deemed that a model is inadequate, new variables that potentially offer better predictive power might be investigated or alternative samples might be examined and a new model developed. It also might be determined through testing or use that a site location model consistently misclassifies certain types of sites. In this case a model designed specifically for that site type might be considered.

The method of sequential analysis (described above) is specifically designed to indicate the need for model revision in an ongoing research framework. When test sampling indicates that model-predicted site densities exceed or fall below specified limits, the model should be rejected. If this should happen the need for model revision is indicated. Even for already tested models, ongoing testing through sequential methods might be conducted as future archaeological surveys are carried out and new information becomes available.

The use of geographic information systems techniques (computer data bases encoded with environmental and other geographic information; see Chapter 10)

might eventually lead to interactive model building, testing, and revision as an ongoing process. If environmental and other variables relevant to archaeological site location models are encoded in the data base along with known site locations, predictive models of many forms and varieties, such as models for multiple site types or temporal periods, could be generated instantaneously. As new sites and nonsites are discovered, additional model tests could be performed or these data could be incorporated into the data base to update existing models. As models change, so do the results of models. Computer graphic techniques can allow new maps of model results to be rapidly and cost-effectively produced so that the most current information can be used.

Many of the results presented here stem from a history of personal involvement in archaeological predictive modeling that spans the past decade. This involvement owes its origins to a 1979 contract awarded by the Bureau of Land Management, Grand Junction District Office, to Nickens and Associates of Montrose, Colorado (where I was employed), for a regional predictive model. The Grand Junction District Office again supported my work in 1982-1983 for a larger study that resulted in considerably improved methods, including extensive survey and model testing, and a procedural manual. Doctoral work at the University of California at Santa Barbara exposed me to geographic information systems and remote sensing technology, the quantitative expertise of Albert C. Spaulding, and the hunter-gatherer settlement ideas of Michael A. Jochim, and it resulted in a dissertation on archaeological predictive modeling in 1983. The University of Denver's Piñon Canyon Archaeological Project, sponsored by the U.S. Army, called for extensive use and application of archaeological predictive models and GIS technology. My involvement with that project from 1983 to 1985 allowed further refinement and development of modeling methods with a very large data set, and production of GIS capabilities compatible with archaeological analysis and modeling needs. In recent years at the University of Arizona my teaching of a spatial analysis class, which focuses on GIS and archaeological modeling, has forced me to express the basic ideas and methods on an easy-to-understand level. Of more importance, however, are the many insights and applications of those technologies that my students have given me. Chapters 7, 8, and 10 of this volume owe much to the above persons and institutions.

If there is any gauge of the success of one's work, it is in how much it is used. Although I have not received royalties as of yet, I am frequently sent copies of project reports that utilize (and in many cases copy directly) the methods summarized in Chapter 8. These studies, largely stemming from cultural resource management contexts and performed for various government agencies, represent a staggering amount of work (probably approaching 100 studies and projects). My hope is that the results of this work will be used responsibly by management personnel as tools to better care for, preserve, and protect cultural resources. If they are not used responsibly, then the fault lies with management (not the models) and we must focus our attention on defining responsibility.

I wish to thank the following individuals in particular for their contributions to Chapters 7, 8, and 10 of this volume. JoAnn Christein devoted considerable effort in manuscript production over several rewrites, in digitization of much of the data for the GIS work, and in giving moral support over this long and trying project. I am grateful to Mike Jochim for allowing me to present some of his Mesolithic data in these chapters. Finally, Dan Martin deserves special praise for his support of the whole volume and particularly for his continued encouragement of my work.

REFERENCES CITED

Adams, E. Charles

1974 Location Strategy Employed by Prehistoric Inhabitants of the Piedra District, Colorado. *Southwestern Lore* 43:1–15.

Anderson, J. A.

1975 Quadratic Logistic Discrimination. *Biometrika* 62:149–154.

Aronoff, Stan

1982 The Map Accuracy Report: A Users View. *Photogrammetric Engineering and Remote Sensing* 48:1309–1312.

Basu, J. P., and P. L. Odell

1974 Effect of Intra Class Correlation Among Training Samples on the Misclassification Probabilities of Bayes Procedure. *Pattern Recognition* 6:13–16.

Baxter, N. D., and J. G. Cragg

1970 Corporate Choice Among Long-Term Financing Instruments. *Review of Economics and Statistics* 6:13–16.

Berry, Kenneth J., Kenneth L. Kvamme, and Paul W. Mielke, Jr.

1980 A Permutation Technique for the Spatial Analysis of the Distribution of Artifact Classes. *American Antiquity* 43:482–486.

1983 Improvements in the Permutation Test for the Spatial Analysis of the Distribution of Artifacts into Classes. *American Antiquity* 48:547–553.

Berry, Kenneth J., Paul W. Mielke, Jr., and Kenneth L. Kvamme

1984 Efficient Permutation Procedures for Analysis of Artifact Class Patterns in Archaeological Space. In *Intrasite Spatial Analysis in Archaeology*, edited by Harold J. Hietala and Paul L. Larson, pp. 54–74. Cambridge University Press, Cambridge, England.

Berry, Michael S.

1984 Sampling and Predictive Modeling on Federal Lands in the West. *American Antiquity* 49:842–853.

Bettinger, Robert L.

1977 Predicting the Archaeological Potential of the Inyo-Mano Region of Eastern California. In *Conservation Archaeology*, edited by Michael B. Schiffer and George J. Gumerman, pp. 217–226. Academic Press, New York.

1980 Explanatory/Predictive Models of Hunter-Gatherer Adaptation. In *Advances in Archaeological Method and Theory*, vol. 3, edited by Michael B. Schiffer, pp. 189–255. Academic Press, New York.

Bove, Frederick J.

1981 Trend Surface Analysis of the Lowland Classic Maya Collapse. *American Antiquity* 46:93–112.

Bradley, John E., William R. Killam, George R. Burns, and Marilyn A. Martorano

1984 *A Class II Survey and Predictive Model of Selected Areas in the Cisco Desert Grand County, Utah.* Goodson and Associates, Inc. Submitted to the Bureau of Land Management. Copies available from Bureau of Land Management, Moab District, Utah.

Brand, R. J., R. H. Rosenman, and R. I. Sholtz

1976 Multivariate Prediction of Coronary Heart Disease in the Western Collaborative Group Study Compared to the Findings of the Framingham Study. *Circulation* 53:348–355.

Brose, David S.

1976 Locational Analysis in the Prehistory of Northeast Ohio. In *Cultural Change and Continuity: Essays in Honor of James Bennett Griffin*, edited by Charles E. Cleland, pp. 3–18. Academic Press, New York.

Brown, Kenneth L.
1979 Late Prehistoric Settlement Patterns in Southwestern Kansas. *Plains Anthropologist* 43:482–486.

Burgess, R. J., K. L. Kvamme, P. R. Nickens, A. D. Reed, and G. C. Tucker, Jr.
1980 *Class II Cultural Resource Inventory of the Glenwood Springs Resource Area, Grand Junction District, Colorado.* Nickens and Associates. Submitted to Bureau of Land Management, Grand Junction District. Copies available from Bureau of Land Management, Grand Junction, Colorado.

Campbell, James B.
1981 Spatial Correlation Effects Upon Accuracy of Supervised Classification of Land Cover. *Photogrammetric Engineering and Remote Sensing* 47:355–363.

Carr, Christopher
1985 Introductory Remarks on Regional Analysis. In *For Concordance in Archaeological Analysis: Bridging Data Structure, Quantitative Technique, and Theory*, edited by Christopher Carr, pp. 114–127. Westporr, Kansas City.

Cliff, A. D., and J. K. Ord
1973 *Spatial Autocorrelation.* Pion, London.

1975 The Comparison of Means When Samples Consist of Spatially Autocorrel ed Observations. *Environment and Planning* A 7:725–734.

1980 *Spatial Processes: Models and Applications.* Pion, London.

Congalton, Russell G., Richard G. Oderwald, and Roy A. Mead
1983 Assessing Landsat Classification Accuracy Using Discrete Multivariate Analysis Statistical Techniques. *Photogrammetric Engineering and Remote Sensing* 49:1671–1678.

Conover, William J.
1971 *Practical Nonparametric Statistics.* John Wiley, New York.

Cooper, R. A., and A. J. Weekes
1983 *Data, Models and Statistical Analysis.* Barnes and Noble, Totowa, New Jersey.

Costanzo, C. M., W. C. Halperin, N. D. Gale, and G. D. Richardson
1982 An Alternative Method for Assessing Goodness-of-Fit for Logit Models. *Environment and Planning* A 14:963–971.

Custer, Jay F., Timothy Eveleigh, and V. Klemas
1983 A Landsat-Generated Predictive Model for Prehistoric Archaeological Sites in Delaware's Coastal Plain. *Bulletin of the Archaeological Society of Delaware* n.s. 14:19–38.

Custer, Jay F., Timothy Eveleigh, V. Klemas, and I. Wells
1986 Application of Landsat Data and Synoptic Remote Sensing to Predictive Models for Prehistoric Sites: An Example from the Delaware Coastal Plain. *American Antiquity* 51:572–588.

Davis, John C.
1973 *Statistics and Data Analysis in Geology.* John Wiley, New York.

Delfiner, P., and J. P. Delhomme
1975 Optimum Interpolation by Kriging. In *Display and Analysis of Spatial Data*, edited by John C. Davis and Michael J. McCullagh, pp. 96–114. John Wiley, New York.

Dixon, Wilfred J., and Frank J. Massey, Jr.
1957 *Introduction to Statistical Analysis.* McGraw-Hill, New York.

Dixon, W. J., M. B. Brown, L. Engelman, J. W. Frane, M. A. Hill, and R. I. Jennrich
1983 *BMDP Statistical Software.* University of California Press, Berkeley.

Duda, R. C., and P. E. Hart
1973 *Pattern Classification and Scene Analysis.* John Wiley, New York.

Dunnell, R. C., and W. S. Dancey
1983 The Siteless Survey: A Regional Scale Data Collection Strategy. In *Advances in Archaeological Method and Theory*, vol. 6, edited by Michael B. Schiffer, pp. 267–287. Academic Press, New York.

Eisenbeis, R. A., and R. B. Avery
1972 *Discriminant Analysis and Classification Procedures.* Lexington Books, Lexington, Massachusetts.

Ericson, Jonathon E., and Raymond Goldstein
1980 Workspace: A New Approach to the Analysis of Energy Expenditure Within Site Catchments. *Anthropology UCLA* 10:21–30.

Euler, Robert C., and Susan M. Chandler
1978 Aspects of Prehistoric Settlement Patterns in Grand Canyon. In *Investigations of the Southwestern Anthropological Research Group: An Experiment in Archaeological Cooperation*, edited by Robert C. Euler and George J. Gumerman. Museum of Northern Arizona, Flagstaff.

Euler, Robert C., and George G. Gumerman, editors
1978 *Investigations of the Southwestern Anthropological Research Group: An Experiment in Archaeological Cooperation.* Museum of Northern Arizona, Flagstaff.

Feder, Kenneth L.
1979 Geographic Patterning of Tool Types as Elicited by Trend Surface Analysis. In *Computer Graphics in Archaeology*, edited by Steadman Upham, pp. 95–102. Anthropological Research Papers No. 15. Arizona State University, Tempe.

Findlow, Frank J.
1980 A Catchment Analysis of San Luis Phase and Animas Phase Sites in Southwestern New Mexico. *Anthropology UCLA* 10:157–178.

Findlow, Frank J., and Jonathan E. Ericson, editors
1980 Catchment Analysis: Essays on Prehistoric Resource Space. *Anthropology UCLA* 10 (whole no.).

Flannery, Kent V.
1972 The Cultural Evolution of Civilizations. *Annual Review of Ecology and Systematics* 3:399–425.

1976 The Village and Its Catchment Area. In *The Early Mesoamerican Village*, edited by Kent V. Flannery, pp. 91–95. Academic Press, New York.

Frank, Ronald E., William F. Massy, and Doland G. Morrison
1965 Bias in Multiple Discriminant Analysis. *Journal of Marketing Research* 2:250–258.

Gould, Peter
1970 Is *Statistix inferens* the Geographical Name for a Wild Goose? *Economic Geography* 46:439–448.

Grady, James
1980 *Environmental Factors in Archaeological Site Location.* Cultural Resource Series 9. Bureau of Land Management, Colorado.

Green, Ernestine L.
1973 Location Analysis of Prehistoric Maya Sites in Northern British Honduras. *American Antiquity* 38:279–293.

Green, Paul E.
1978 *Mathematical Tools for Applied Multivariate Analysis*, 2nd ed. Academic Press, New York.

Grossman, L.
1977 Man-Environment Relationships in Anthropology and Geography. *Annals of the Association of American Geographers* 67:127–144.

Gumerman, George J., editor
1971 *The Distribution of Prehistoric Population Aggregates.* Prescott College Anthropological Reports No. 1. Prescott, Arizona.

Gumerman, George J., and R. Roy Johnson
1971 Prehistoric Human Population Distribution in a Biological Transition Zone. In *The Distribution of Prehistoric Population Aggregates*, edited by George J. Gumerman, pp. 83-102. Prescott College Anthropological Reports No. 1. Prescott, Arizona.

Haggett, P., A. D. Cliff, and A. Frey
1977 *Locational Analysis in Human Geography*, 2nd ed. John Wiley, New York.

Haining, R. P.
1981 Spatial and Temporal Analysis: Spatial Modelling. In *Quantitative Geography: A British View*, edited by N. Wrigley and R. J. Bennett, pp. 86-91. Routledge and Kegan Paul, London.

Hammond, Edwin H.
1964 Analysis of Properties of Land Form Geography: An Application to Broad-Scale Land Form Mapping. *Annals of the Association of American Geographers* 54:11-19.

Hartung, R. E., and W. J. Lloyd
1969 Influence of Aspects on Forests of the Clarksville Soils in Dent County, Missouri. *Journal of Forestry* 67:178-182.

Hasenstab, Robert J.
1983 The Application of Geographic Information Systems to the Analysis of Archaeological Site Distributions. Paper presented at the 48th Annual Meeting of the Society for American Archaeology, Pittsburgh.

Hay, Alan M.
1979 Sampling Designs to Test Land-Use Map Accuracy. *Photogrammetric Engineering and Remote Sensing* 45:529-533.

Hays, William L.
1981 *Statistics*, 3rd ed. Holt, Rinehart and Winston, New York.

Hietala, Harold J., and Paul L. Larson
1979 SYMAP Analysis in Archaeology: Intrasite Assumptions and a Comparison with TREND Analysis. *Norwegian Archaeological Review* 12:57-64.

Hill, James N.
1971 Research Propositions for Consideration: Southwestern Anthropological Research Group. In *The Distribution of Prehistoric Population Aggregates*, edited by George J. Gumerman, pp. 55-62. Prescott College Anthropological Reports No. 1. Prescott, Arizona.

Hixon, M. M., B. J. Davis, and M. G. Bauer
1980 Sampling Landsat Classification for Crop Area Estimations. *Photogrammetric Engineering and Remote Sensing*, pp. 1343-1348.

Hodder, Ian, and Clive Orton
1976 *Spatial Analysis in Archaeology*. Cambridge University Press, Cambridge, England.

Holmer, Richard N.
1979 *Split Mountain Cultural Study Tract*. Utah Archaeological Center. Submitted to the Bureau of Land Management. Copies available from Bureau of Land Management, Vernal District, Utah.

1982 *Castle Valley Rail Line, Denver and Rio Grande Western Railroad: Known Archaeological and Historical Site Evaluation and Predictive Model*. University of Utah Archaeological Center. Submitted to Sverdrup and Parcel & Associates, Inc. Copies available from University of Utah Archaeological Center.

Holmes, John H.
1970 The Theory of Plane Sampling and Its Applications in Geographic Research. *Economic Geography* 46:379-392.

Hord, R. Michael, and William Brooner
1976 Land-Use Map Accuracy Criteria. *Photogrammetric Engineering and Remote Sensing* 42:671-677.

Hudson, J. C.
1969 A Location Theory for Rural Settlement. *Annals of the Association of American Geographers* 59:365–381.

Hurlbett, Robert E.
1977 *Environmental Constraint and Settlement Predictability, Northwestern Colorado.* Cultural Resource Series 3. Bureau of Land Management, Denver, Colorado.

Jochim, Michael A.
1976 *Hunter-Gatherer Subsistence and Settlement: A Predictive Model.* Academic Press, New York.

Johnson, Gregory A.
1972 A Test of the Utility of Central Place Theory in Archaeology. In *Man, Settlement and Urbanism*, edited by Peter J. Ucko, Ruth Tringham, and G. W. Dimbleby, pp. 769–785. Duckworth, England.

1977 Aspects of Regional Analysis in Archaeology. *Annual Review of Anthropology* 6:479–508.

Judge, W. James
1973 *Paleoindian Occupation of the Central Rio Grande Valley, New Mexico.* University of New Mexico Press, Albuquerque.

Kohler, Timothy A., and Sandra C. Parker
1986 Predictive Models for Archaeological Resource Location. In *Advances in Archaeological Method and Theory*, vol. 9, edited by Michael B. Schiffer, pp. 397–452. Academic Press, New York.

Kvamme, Kenneth L.
1980 Predictive Model of Site Location in the Glenwood Springs Resource Area. In *Class II Cultural Resource Inventory of the Glenwood Springs Resource Area, Grand Junction District, Colorado*, Part II, by Robert J. Burgess et al., pp. 88–136. Nickens and Associates. Submitted to Bureau of Land Management. Copies available from Bureau of Land Management, Grand Junction District, Colorado.

1983a *New Methods for Investigating the Environmental Basis of Prehistoric Site Locations.* Unpublished Ph.D. dissertation, Department of Anthropology, University of California, Santa Barbara.

1983b Computer Processing Techniques for Regional Modeling of Archaeological Site Locations. *Advances in Computer Archaeology* 1:26–52.

1983c A Manual for Predictive Site Location Models: Examples from the Grand Junction District, Colorado. Draft submitted to the Bureau of Land Management, Grand Junction District, Colorado. Copies available from Bureau of Land Management, Grand Junction District, Colorado.

1984 Models of Prehistoric Site Location Near Pinyon Canyon, Colorado. In *Papers of the Philmont Conference on the Archaeology of Northeastern New Mexico*, edited by Carol J. Condie, pp. 347–370. Proceedings of the New Mexico Archaeological Council 6(1). Albuquerque.

1985a Determining Empirical Relationships between the Natural Environment and Prehistoric Site Locations: A Hunter-Gatherer Example. In *For Concordance in Archaeological Analysis: Bridging Data Structure, Quantitative Technique, and Theory*, edited by Christopher Carr, pp 208–238. Westport Press, Kansas City.

1985b Geographic Information Systems Techniques for Regional Archaeological Research. Paper presented at the International Symposium on Data Management and Mathematical Models in Archaeology, International Union of Pre- and Protohistoric Sciences, Denver.

1986 The Use of Geographic Information Systems for Modeling Archaeological Site Distributions. In *Geographic Information Systems in Government*, vol. I, edited by B. K. Opitz, pp. 345–362. A. Deepak, Hampton, Virginia.

Kvamme, Kenneth L., and Michael A. Jochim
1988 Environmental Basis of Mesolithic Settlement. In *The Mesolithic in Europe: Papers Presented at the Third International Symposium*, edited by Clive Bonsall. John Donald, Edinburgh, Scotland, in press.

Lachenbruch, P. A., and M. Mickey
1968 Estimation of Error Rates in Discriminant Analysis. *Technometrics* 10:1–11.

Lafferty, Robert H., III
1981 Distribution of Archaeological Materials. In *Settlement Predictions in Sparta*, by Robert H. Lafferty III, Jeffrey L. Otinger, Sandra C. Scholtz, W. Frederick Limp, Beverly Watkins, and Robert D. Jones, pp. 163–206. Arkansas Archaeological Survey Research Series No. 14.

Lafferty, Robert H., III, Jeffrey L. Otinger, Sandra C. Scholtz, W. Frederick Limp, Beverly Watkins, and Robert D. Jones
1981 *Settlement Predictions in Sparta*. Arkansas Archaeological Survey Research Series No. 14.

Landgrebe, David A.
1978 The Quantitative Approach: Concepts and Rationale. In *Remote Sensing: The Quantitative Approach*, edited by Peter H. Swain and Shirley M. Davis. McGraw-Hill, New York.

Larralde, Signa L., and Susan M. Chandler
1981 *Archaeological Inventory in the Seep Ridge Cultural Study Tract, Uinta County, Northeastern Utah*. Cultural Resource Series 11. Bureau of Land Management, Salt Lake City.

Larson, Paul
1975 *Trend Analysis in Archaeology: A Preliminary Study of Intrasite Patterning*. Norwegian Archaeological Review 8:75–80.

Lemeshow, Stanley, and David W. Hosmer, Jr.
1982 A Review of Goodness-of-Fit Statistics for Use in the Development of Logistic Regression Models. *American Journal of Epidemiology* 115:92–106.

Limp, Frederick, and Robert M. Lafferty
1981 Conclusions and Recommendations. In *Settlement Predictions in Sparta*, by R. H. Lafferty III, J. L. Otinger, S. C. Scholtz, W. F. Limp, B. Watkins, and R. D. Jones, pp. 223–240. Arkansas Archaeological Survey Research Series No. 14.

Lin, G. C., and T. C. Minter
1976 Bayes Estimation on Parameters of the Single-Class Classifier. In *Symposium on Machine Processing of Remotely Sensed Data*, pp. 22–28. The Institute of Electrical and Electronics Engineers, Purdue University, Lafayette, Indiana.

Lipe, W. D., and R. G. Matson
1971 Human Settlement and Resources in the Cedar Mesa Area, Southeastern Utah. In *The Distribution of Prehistoric Population Aggregates*, edited by George J. Gumerman, pp. 126–151. Prescott College Anthropological Reports No. 1. Prescott, Arizona.

Lovis, W. A., Jr.
1976 Quarter Sections and Forests: An Example of Probability Sampling in the Northeastern Woodlands. *American Antiquity* 41:364–372.

Maynard, Paul F.
1981 *The Logit Classifier: A General Maximum Likelihood Discriminant for Remote Sensing Applications*. Unpublished M.A. thesis, Department of Geography, University of California, Santa Barbara.

Maynard, Paul F., and Alan H. Strahler
1981 Logit Classifier, a General Maximum Likelihood Discriminant for Remote Sensing Applications. *Proceedings of the Fifteenth International Symposium on Remote Sensing of Environment*, pp. 213–222. Ann Arbor.

Michaelis, J.
1973 Simulation Experiments with Multiple Group Linear and Quadratic Discriminant Analysis. In *Discriminant Analysis and Applications*, edited by T. Cacoullos, pp. 225–238. Academic Press, New York.

Miller, R. G.
1966 *Simultaneous Statistical Inference*. McGraw-Hill, New York.

Minter, T. C.
1975 The Single-Class Maximum Likelihood Classifier. In *Symposium on Machine Classification of Remotely Sensed Data*. The Institute of Electrical and Electronics Engineers, Purdue University, Lafayette, Indiana.

Missalati, A., A. E. Prelat, and R. J. P. Lyon
1979 Simultaneous Use of Geological, Geophysical, and Landsat Digital Data in Uranium Exploration. *Remote Sensing of Environment* 8:189–210.

Moik, Johannes G.
1980 *Digital Processing of Remotely Sensed Images*. National Aeronautics and Space Administration. Government Printing Office, Washington, D.C.

Monroe, C. B., K. L. Feder, D. A. Poirier, and R. R. Gradie
1980 Trend Surface Analysis: The Settlement Dynamics of Colonial Connecticut. *Man in the Northeast* 19:105–128.

Morain, S. A., C. Nelson, M. E. White, and A. M. Komorek
1981 Remote Detection of Prehistoric Sites in Bandelier National Monument. In *Remote Sensing Multispectral Analysis of Cultural Resources: Chaco Canyon and Bandelier National Monument*, edited by T. R. Lyons. Remote Sensing: A Handbook for Archeologists and Cultural Resource Managers, Supplement No. 5. National Park Service, Washington, D.C.

Morrison, Donald F.
1976 *Multivariate Statistical Methods*, 2nd ed. McGraw-Hill, New York.

Morrison, Donald G.
1969 On the Interpretation of Discriminant Analysis. *Journal of Marketing Research* 6:156–163.

Mosteller, Frederick, and John W. Tukey
1977 *Data Analysis and Regression*. Addison-Wesley, Reading, Massachusetts.

Mueller, James W.
1974 *The Use of Sampling in Archaeological Survey*. Society for American Archaeology Memoir 28.

Mueller, James W., editor
1975 *Sampling in Archaeology*. University of Arizona Press, Tucson.

Overall, J. E., and G. J. Klett
1972 *Applied Multivariate Analysis*. McGraw-Hill, New York.

Parker, Sandra C.
1985 Predictive Modeling of Site Settlement Systems Using Multivariate Logistics. In *For Concordance in Archaeological Analysis: Bridging Data Structure, Quantitative Techniques, and Theory*, edited by Christopher C. Carr, pp. 173–207. Westport, Kansas City.

Peebles, Thomas C., editor
1981 *A Class II Cultural Resources Survey of the Eastern Powder River Basin, Wyoming*. Metcalf-Zier Archaeologists, Inc. Submitted to Bureau of Land Management. Copies available from Bureau of Land Management, Wyoming State Office, Cheyenne.

Pindyck, Robert S., and Daniel L. Rubinfeld
1976 *Econometric Models and Economic Forecasts*. McGraw-Hill, New York.

Plog, Fred T.
1968 *Archaeological Surveys: A New Perspective.* Unpublished M.A. thesis, Department of Anthropology, University of Chicago.

1971 Some Operational Considerations. In *The Distribution of Prehistoric Population Aggregates,* edited by George J. Gumerman, pp. 45–54. Prescott College Anthropological Reports No. 1. Prescott, Arizona.

1981 SARG: The Computer in a Cooperative Effort. In *Data Bank Applications in Archaeology,* edited by Sylvia W. Gaines, pp. 46–56. University of Arizona Press, Tucson.

1983 Study Area 1: Kaibab National Forest. In *Theory and Model Building: Refining Survey Strategies for Locating Prehistoric Heritage Resources,* edited by Linda S. Cordell and Dee F. Green, pp. 63–66. Cultural Resources Document No. 3. USDA Forest Service.

Plog, Fred T., and James N. Hill
1971 Explaining Variability in the Distribution of Prehistoric Population Aggregates. In *The Distribution of Prehistoric Population Aggregates,* edited by George J. Gumerman, pp. 7–36. Prescott College Anthropological Reports No. 1. Prescott, Arizona.

Press, S. J., and S. Wilson
1979 Choosing Between Logistic and Regression Discriminant Analysis. *Journal of the American Statistical Association* 73:699–705.

Reed, Alan D., and Susan M. Chandler
1984 *A Sample-Oriented Cultural Resource Inventory in Carbon, Emery, and San Pete Counties, Utah.* Nickens and Associates. Submitted to Bureau of Land Management. Copies available from Bureau of Land Management, Salt Lake City.

Reher, Charles A.
1977 Settlement and Subsisrence along the Lower Chaco River. In *Settlement and Subsistence along the Lower Chaco River,* edited by Charles Reher, pp. 7–112. University of New Mexico Press, Albuquerque.

Robinove, Charles J.
1981 The Logic of Multispectral Classification and Mapping of Land. *Remote Sensing of Environment* 11:231–244.

Roper, Donna C.
1976 A Trend-Surface Analysis of Central Plains Radiocarbon Dates. *American Antiquity* 41:181–188.

1979a The Method and Theory of Site Catchment Analysis: A Review. In *Advances in Archaeological Method and Theory,* vol. 2, edited by Michael B. Schiffer, pp. 120–140. Academic Press, New York.

1979b *Archaeological Survey and Settlement Pattern Models in Central Illinois.* Scientific Papers 16. Illinois State Museum, Springfield.

Rosenfield, George H.
1981 Analysis of Variance of Thematic Mapping Experiment Data. *Photogrammetric Engineering and Remote Sensing* 47:1685–1692.

Rosenfield, G. H., K. Fitzpatrick-Lins, and H. S. Ling
1982 Sampling for Thematic Map Accuracy Testing. *Photogrammetric Engineering and Remote Sensing* 48:131–137.

SAS Institute
1982 *SAS User's Guide: Statistics.* SAS Institute, Cary, North Carolina.

Schmidt, Peter, and Robert P. Strauss
1975 The Prediction of Occupation Using Multiple Logit Models. *International Economic Review* 16:471–486.

Scholtz, Sandra C.
1981 Location Choice Models in Sparta. In *Settlement Predictions in Sparta*, by R. Lafferty III, J. L. Otinger, S. C Scholtz, W. F. Limp, B. Watkins, and R. O. Jones, pp. 207–222. Arkansas Archaeological Survey Research Series No. 14.

Schowengerdt, Robert A.
1983 *Techniques for Image Processing and Classification in Remote Sensing.* Academic Press, New York.

Schroedl, Alan R.
1984 Site Location Prediction and Modeling. In *The Tar Sands Project: Cultural Resource Inventory and Predictive Modeling in Central and Southern Utah*, by Betsy L. Tipps. P-III Associates. Submitted to the Bureau of Land Management. Copies available from Bureau of Land Management, Salt Lake City.

Schuldenrein, Joseph
1983 Pinyon Canyon Interim Management Report. Commonwealth Associates. Submitted to National Park Service, Interagency Archeological Services, Denver. Copies available from IAS, Denver.

Smith, C. A. B.
1947 Some Examples of Discrimination. *Annals of Eugenics* 13:272–282.

Steel, R. G. D., and J. H. Torrie
1980 *Principles and Procedures of Statistics*, 2nd ed. McGraw-Hill, New York.

Stopher, P. R., and A. H. Meyburg
1979 *Survey Sampling and Multivariate Analysis for Social Scientists and Engineers.* Lexington Books, Lexington, Massachusetts.

Swain, Phillip H.
1978 Fundamentals of Pattern Recognition in Remote Sensing. In *Remote Sensing: The Quantitative Approach*, edited by Phillip H. Swain and Shirley M. Davis, pp. 136–187. McGraw-Hill, New York.

Swain, Phillip H., and Shirley M. Davis, editors
1978 *Remote Sensing: The Quantitative Approach.* McGraw-Hill, New York.

Thomas, David H.
1975 Nonsite Sampling in Archaeology: Up the Creek Without a Site? In *Sampling in Archaeology*, edited by James W. Mueller, pp. 61–81. University of Arizona Press, Tucson.

1976 *Figuring Anthropology: First Principles of Probability and Statistics.* Holt, Rinehart and Winston, New York.

1979 *Archaeology.* Holt, Rinehart and Winston, New York.

Thomas, David H., and Robert L. Bettinger
1976 *Prehistoric Piñon Ecotone Settlements of the Upper Reese River Valley, Central Nevada.* Anthropological Papers of the American Museum of Natural History 53(3). New York.

Thompson, R. A.
1978 *A Stratified Random Sample of the Cultural Resources in the Canyonlands Section of the Moab District of the Bureau of Land Management in Eastern Utah.* Cultural Resources Series No. 1. Bureau of Land Management, Utah.

Tobler, Waldo R.
1970 A Computer Movie Simulating Urban Growth in the Detroit Region. *Economic Geography* 46:234–240.

Tubbs, J. D., and W. A. Coberly
1978 Spatial Correlation and Its Effect Upon Classification in Landsat. *Proceedings of the Twelfth International Symposium on Remote Sensing of Environment*, pp. 775–781. Ann Arbor.

Unwin, David J.
1975 *An Introduction to Trend Surface Analysis.* Concepts and Techniques in Modern Geography No. 5. GeoAbstracts, University of East Anglia, Norwich.

1981 *Introductory Spatial Analysis.* Methuen, London.

Van Ness, Margaret A.
1984 Analysis of Vegetation Patterns at Piñon Canyon, Colorado. Archaeological Research Institute, University of Denver. Submitted to National Park Service, Interagency Archeological Services, Denver. Copies available from IAS, Denver.

Vivian, R. Gwinn, L. Vogler, A. Sullivan, and A. Rieger
1980 AZSITE: A Computerized SELGEM File for the Arizona State Museum Site Survey Records. In *The Arizona State Museum Archaeological Site Survey System,* compiled by Lawrence E. Vogler, pp. 41–102. Archaeological Series No. 128. Arizona State Museum, Tucson.

Wells, L., J. Custer, and V. Klemas
1981 Locating Prehistoric Archaeological Sites Using Landsat. *Proceedings of the Fifteenth International Symposium on Remote Sensing of Environment.* Ann Arbor.

Western, D., and T. Dunne
1979 Environmental Aspects of Settlement Site Decisions Among Pastoral Maasai. *Human Ecology* 7:75–98.

Wetherill, G. Barrie
1975 *Sequential Methods in Statistics.* Chapman and Hall, London.

Williams, L., D. H. Thomas, and R. L. Bettinger
1973 Notions to Numbers: Great Basin Settlements as Polythetic Sets. In *Research and Theory in Current Archaeology,* edited by C. L. Redman, pp. 215–238. John Wiley, New York.

Wilmsen, Edwin N.
1973 Interaction, Spacing Behavior, and the Organization of Hunting Bands. *Journal of Anthropological Research* 29:1–31.

Wobst, H. Martin
1976 Locational Relationships in Paleolithic Society. *Journal of Human Evolution* 5:49–58.

Wood, J. J.
1978 Optimal Locations in Settlement Space: A Model for Describing Location Strategies. *American Antiquity* 43:258–270.

Wormington, H. M., and Robert H. Lister
1956 *Archaeological Investigations on the Uncompahgre Plateau.* Denver Museum of Natural History Proceedings No. 2.

Wrigley, Neil
1976 *An Introduction to the Use of Logit Models in Geography.* Concepts and Techniques in Modern Geography No. 10. GeoAbstracts, University of East Anglia, Norwich.

1977a Probability Surface Mapping: A New Approach to Trend Surface Mapping. *Transactions of the Institute of British Geographers* 2:129–140.

1977b *Probability Surface Mapping.* Concepts and Techniques in Modern Geography No. 16. GeoAbstracts, University of East Anglia, Norwich.

1982 Quantitative Methods: Developments in Discrete Choice Modeling. *Progress in Human Geography* 6:547–562.

Zarky, A.
1976 Statistical Analysis of Site Catchments at Ocos, Guatemala. In *The Early Mesoamerican Village,* edited by Kent V. Flannery, pp. 117–130. Academic Press, New York.

Zier, Anne H., and T. C. Peebles
1982 *Report on the Kemmerer Resource Area Class II Cultural Resource Inventory, Lincoln and Uinta Counties, Wyoming*. Metcalf-Zier Archaeologists, Inc. Submitted to Bureau of Land Management. Copies available from Bureau of Land Management, Rock Springs District, Wyoming.

Zubrow, E. B., and J. W. Harbaugh
1978 Archaeological Prospecting: Kriging and Simulation. In *Simulation Studies in Archaeology*, edited by Ian Hodder, pp. 109–122. Cambridge Press, Cambridge, England.

Chapter 9

REMOTE SENSING IN ARCHAEOLOGICAL PROJECTION AND PREDICTION

James I. Ebert

During the preparatory stages of this volume the authors and editors were not certain that a chapter on the potential and use of remote sensing in archaeological predictive modeling would be entirely appropriate. The purposes of this book are to explore some of the complexities of predictive modeling, to examine some of the biases inherent in our present methods and data (see particularly Chapter 7), and to suggest directions that archaeological explanation will have to take in order to achieve successful and scientifically useful predictions (Chapter 4). There was some concern that the inclusion of a chapter on using remote sensing to do predictive modeling might imply that technical means now exist by which predictions can easily be made, that all one has to do is plug existing archaeological and remote-sensing-derived environmental data into a computer and a predictive model will emerge.

It is clear from the preceding chapters in this volume that this is not the case. Predictive modeling is an area of great interest to archaeologists and managers alike, and perhaps more than any other fact, this interest indicates that we are just beginning to understand how to predict and model. One of the most universal cultural patterns is that people worry about and try to predict things in inverse proportion to how well they can really predict them. Nightly weather forecasts, for instance, dwell heavily on such questions as whether it will rain tomorrow, and the resultant predictions are of mixed success at best; there is never any discussion about whether the sun will come up in the morning. When we finally do perfect archaeological predictive modeling, there will probably be little discussion about it at meetings or in the literature. As discussed in Chapter 4, however, before we achieve success in prediction we will have had to learn many other things—how human systems are organized at several levels; how deposition and postdepositional processes affect the preservation and visibility of archaeological materials, and how this varies across the landscape; and how to make our methods of data discovery, collection, and analysis compatible with what we want to know about the past. In short, by the time we know how to do prediction we will also have discovered how to explain the archaeological record, and by the time we know how to predict, we may not need to do so anymore.

At present, we have not achieved any of these goals completely. The preceding chapters of this volume point out that there is still a great deal of archaeological research to be done. There is not, in fact, agreement within the profession even about what predictive modeling means or about definitions of such important basic operational terms as *site* or *system*. This may seem unfortunate, but it need not be thought of as being so. Learning how to do predictive modeling and archaeology in general is a great adventure that we have just begun.

This chapter will explore the possible role that remote sensing can play in that adventure and describe attempts by archaeologists to use remote sensing to project, predict, and explain the archaeological record, the operation of past behavior and behavioral systems, and the things that separate these two domains. This chapter will begin with a review of the basics of remote sensing—what it is, the methods and techniques by which it is carried out, the data that it yields, and its capabilities and limitations for archaeological projection and prediction. Relevant literature and contemporary attempts at incorporating remote sensing in archaeological projection and prediction will be surveyed, and the strengths and weaknesses of these approaches discussed. Finally, some suggestions will be made about new, potentially productive applications of predictive remote sensing.

FUNDAMENTALS OF REMOTE SENSING

Platforms, Recording Devices, Data Types, and Analyses

Remote sensing is the science and technology of obtaining information or data about physical objects and the environment through the process of recording, measuring, and interpreting photographic images and patterns of electromagnetic radiant energy (Ebert 1984:293). The most familiar remote sensing methods are photographic, and aerial and ground-based photography has been employed in archaeology since the beginnings of the discipline. The term *remote sensing* was coined in the late 1960s in response to the need for a term that could include both simple photographic data-collection techniques and the use of other, more exotic data sources, such as satellite and airborne multispectral scanners and microwave (radar) sensors, in a unified technical field.

Remote sensing can best be understood when broken down into several of its component parts. Remote sensing platforms (the vantage points from which data are collected) range from the surface of the earth to low-altitude camera supports, such as bipods and tripods, to balloons, aircraft, and satellites hundreds of miles above the landscape. The devices with which remote sensor data are collected include active radar transmitters and receivers, proton magnetometers, cameras, and scanning devices recording reflected radiation. Remote sensor data can be recorded by these devices photochemically (i.e., with photographic emulsions) or electronically in either analog or digital formats, and in one or more wide or restricted wavelength bands.

Remote sensor data can be analyzed by human interpreters who *(a)* simply look at visual products or *(b)* use magnifying and projecting devices to examine minute details of an image or *(c)* employ stereoscopes, which produce a three-dimensional image from partially overlapping photographic prints. Stereoscopic images can also be used to produce orthophotos (photographs in which errors of scale and orthographic errors have been removed) and photogrammetric maps, the most familiar of which are USGS topographic maps, virtually all of which are made from aerial photographs.

The advent of computers capable of digital image processing has made available new and versatile forms of remote sensor data analysis based on mathematical manipulation of the matrix of picture elements *(pixels)* that constitute a digital image. Each pixel making up a digital image has a numerical value, which expresses the reflectance of the represented portion of the earth's surface. These values can be subjected to filtering, classification, histogram stretching (contrast enhancement), density slicing (density range simplification), power spectrum analysis, geometric correction, resampling, pattern recognition routines, and virtually any other matrix operation. While digital data are directly derived by most scanning devices, photographic and other analog data can be digitized into pixels for digital analysis, and conversely, digital data can be converted into visual images for photointerpretation.

Clearly, remote sensing encompasses a great many methods and techniques; it is beyond the scope of this chapter to describe and explain each of them. The fundamentals and details of remote sensor platforms, data collection devices, data types, and data analysis devices and methods are covered exhaustively in many available sources to which the reader should refer for more complete information. One of the most comprehensive of these sources is the American Society of Photogrammetry's *Manual of Remote Sensing* (Colwell 1983); one chapter in that volume (Ebert and Lyons 1983) focuses on archaeological, anthropological, and cultural resource remote sensing. Other excellent general remote sensing references are Avery (1977) and Lillesand and Kiefer (1979). A more concise summary of general archaeological applications of remote sensing can be found in Ebert (1984).

Scales and Resolution

Regardless of data source or type, there are two basic properties shared by all remote sensor data: *scale* and *resolution*. The scale of an image refers to the relationship between the size of the image and the actual size of the scene that the image represents. The scale of an image is determined by the distance between the data collection device and the scene being imaged and by the field of view of the data collection device. For aerial photographic data, for instance, the scale equals the focal length of the lens divided by flight height. The scale is generally expressed as a ratio of 1:*x* (distance on the photograph:actual scene distance; Avery 1977:43). As an example, in a photograph with a scale of 1:12,000, 1 cm on the photographic image

would represent a ground distance of 12,000 cm or 120 m. The scale of data recorded by such digital devices as multispectral scanners is determined similarly by distance (altitude) and the instantaneous field of view of the scanner (Lillesand and Kiefer 1979:396).

Remote sensor data resolution is a more complex concept than scale because it can be of three basic sorts: spatial, radiometric, and temporal. Spatial resolution refers to the minimum size of actual objects that can be discerned in an image. This varies with recording medium parameters (photographic emulsions have the highest resolution of any remote sensing recording medium) but in all cases is also a direct function of image scale. The smaller the scale of an image (that is, the smaller the fraction of image size:object size), the lower the resolution. Since larger scale images cover a smaller area than smaller scale images, there is always an economic trade-off between scale and spatial resolution in remote sensing.

Radiometric resolution refers to the portion or portions of the electromagnetic spectrum recorded in remote sensor data. Panchromatic photographs record the same portion of the electromagnetic spectrum seen by the human eye; other photographic emulsions record ultraviolet or near infrared radiation. Microwave (radar) devices record wavelengths much longer than ultraviolet light, while scanners can record visual through far-infrared spectra. Film/filter combinations can restrict the portions of the spectrum that cameras measure, and multiband camera clusters have been used to produce multispectral photographic data. Multispectral scanners (MSS) record more than one wavelength band; the example of multispectral scanner data most familiar to and most frequently used by archaeologists is Landsat, which is discussed at greater length below.

Temporal resolution is a measure of how frequently a scene is imaged through repeated aerial photographic overflights or satellite sensor passes. Comparison of aerial photographs from the 1930s with those taken more recently provides one example of temporal resolution, but the term takes on a clearer meaning in reference to regularly repeated satellite data collection. The Landsat satellites, for instance, cover the entire surface of the earth (cloud conditions permitting) about every 18 days. Temporal resolution is important because the surface and near surface of the earth changes on both large time scales (e.g., geological and geomorphological change) and small time scales (e.g., seasonal variation, vegetational change, and modern development), and change at either scale may be important archaeologically.

Remote Sensor Data for Projection and Prediction

Most archaeological uses of remote sensing that can be characterized as projective or predictive make use of two general data sources: aerial photographs and airborne or satellite multispectral scanner products. Archaeologists have made use of both existing data and data acquired specifically for their projects.

Existing remote sensor data are of course the least expensive to use, for the costs of their acquisition have already been met by others, and copies of the results can be obtained cheaply. Aerial photographs at a vatiety of scales and in several emulsions—black and white, color, and black-and-white or color infrared—have been taken of almost all the continental United States and much of the world. Many aerial photographs are available from government agencies at nominal cost (for a list of these see May 1978 or Ebert 1984). The earliest systematic vertical aerial photographic coverage of parts of the United States was initiated in the 1930s by the Soil Conservation Service, Department of Agriculture, and beginning at about the same time the U.S. Geological Survey began blanket coverage of the country for topographic mapping purposes. Since that time, other government agencies have been taking aerial photographs at an ever-increasing rate. Generally at least one and often five or ten different types of aerial photographs will be available for a given area of interest. For many purposes and in many project areas, it is likely that existing aerial photographs will meet at least some of the archaeological or managerial remote sensing needs. It is also likely, however, that no existing aerial photographs will satisfy all perceived needs. Most government agency aerial mapping photographs are taken at scales smaller than 1:15,000 (1 cm on the photo represents 150 m on the ground), and many are at very small scales, up to 1:400,000.

Sometimes it may be necessary to acquire new, project-specific aerial photographs to meet certain scale and resolution needs. It may also be the case that the time of day or year in which existing photos were taken, or their emulsions, leave something to be desired. Flying new photos may at first appear to be an expensive solution, and certainly it is more expensive than buying photographic prints from the USGS or other agencies. Effectiveness must also be considered, however, and often this concern may outweigh high acquisition costs, especially if no other suitable photographs are available (Avery and Lyons 1981:18).

The other remote sensor data source commonly employed in archaeological efforts toward projection and prediction consists of images derived from the digital multispectral scanners aboard the Landsat satellites. The first Landsat (then called ERTS-I) was launched by the USGS in 1972; since that time four Landsat satellites have been launched and have provided millions of images of the earth's surface. Landsats 1 and 2 orbit the earth 14 times a day in a circular orbit about 900 km above the earth; each covers a 185 km swath with little side-to-side overlap at the equator and as much as 85 percent at 81° north and south latitude. The satellites' orbits are sun-synchronous, and images are always collected at mid-morning. Landsats 1 and 2 collect data in four bands designed to provide a contrasting basis for discriminating between water and land and among different sorts of vegetation cover and different surficial deposits. Landsat data are resampled and corrected after being sent to earth, and the resulting resolution of Landsat 1 and 2 data is 80 by 80 m pixels. Landsat 3 has the same radiometric resolution in four bands, with the addition of a thermal infrared band and a somewhat higher (55 by 55 m) resolution. Landsat 4, launched in 1984, has seven spectral bands and even greater resolution. For a detailed discussion of the parameters of the Landsat satellite sensor systems and their products, see Lillesand and Kiefer (1979:530–583).

The Landsat satellites are designed to provide versatile data with medium spatial resolution, high temporal resolution, and radiometric resolution that would make their data ideal for earth resources studies and assessments. Even with a 55 by 55 m pixel spatial resolution, it is clear that very few archaeological sites or materials will be visible on Landsat MSS data. Landsat data are, however, ideal for analyzing, measuring, and mapping what archaeologists think of as "independent variables," whether these be assumed landscape preferences of past people, ecosystemic variables affecting the placement of human systems and their components, or depositional and postdepositional processes that affect the preservation or visibility of the archaeological record. It is small wonder that archaeologists interested in predicting have made use of Landsat in a variety of ways, and it is likely that Landsat data and perhaps data from similar, soon-to-be-launched satellite sensors (the French SPOT, for instance) will constitute a major resource for such experiments.

Remote Sensor Data Analysis Methods and Techniques

Two basic interpretive or analytical methods have been used by archaeologists who have incorporated remote sensing in their projective and predictive experiments. The first of these is visual interpretation. As noted above, visual interpretation is accomplished by looking at an image in one of several ways. Aerial photographs or visual images derived from other analog or digital remote sensing sources, such as Landsat, can simply be inspected without optical aids, with images being viewed either singly or overlaid in mosaic form. The interpreter, making use of internalized knowledge about how certain landforms or other characteristics of the environment should appear, makes judgments about areas or zones of differential occurrence of these characteristics on the basis of photographic "cues," including tone, color, texture, pattern, shape, and relationship of one photographically imaged feature to another (for a more complete discussion of these image properties, see Ray 1980:6–13).

In stereoscopic photointerpretation, an interpreter views two partially overlapping photographs, each taken from a different position along a flight line; this is usually accomplished with the aid of a stereoscope, which allows the viewer to see one photograph with each eye. This results in the perception of a three-dimensional image in which the vertical dimension is exaggerated because of the wide spacing of the points from which the stereo photos were taken relative to the spacing between human eyes. Small topographic differences are thus easily distinguished, giving clues to landform and the identity or nature of other characteristics of the scene viewed; Ray (1980:14) estimates that topographic differences of as little as 1 ft can be discerned by the average interpreter using a stereoscope and 1:20,000 scale aerial photos.

Photointerpretation might be thought of as being subjective, and to a certain extent it is. Human interpreters, especially those with experience in photointerpretation, possess extensive internalized information about what different landscape

features and other environmental characteristics "should" look like and may stretch these interpretations or generalize boundaries. On the other hand, this internal information allows an interpreter to make supported guesses—usually correct ones—about phenomena not previously experienced. This is something that even the most sophisticated image-processing machines cannot do and is the reason that image analysis cannot, at least at present, be totally automatic.

It should be noted at this point that all map making is a process of interpretation. Most topographic maps in use today, including the USGS topographic maps used in many experiments in archaeological projection and prediction, are compiled using aerial photographs as a primary or exclusive data source. The topographic contours are measured and drawn from the three-dimensional data contained in vertical-axis, stereo aerial photographs using optical-mechanical or analytical photogrammetric plotting devices. While this process is, to a certain extent, subjective, it is quite accurate and precisely repeatable. The indicated degree of slope may be less so, however, as contour lines too close together to be separated during printing are often artificially spread apart. Almost all the rest of the data shown on topographic maps are subjectively interpreted and generalized—including the intermittency and even the existence of water in streams or springs, and the boundaries of forested vs nonforested lands. Maps are interpretations, and when using them for a specific purpose one must ask what the purpose of the interpreter was. For this reason it may well be best to rely on one's own "first generation" interpretation from aerial photographs rather than on the standardized subjectivity of USGS maps for measurement of landform and environmental variables.

The second class of methods used by archaeologists in analyzing remote sensor data for projective or predictive purposes is encompassed by digital analysis. Digital analysis is done by subjecting a matrix of pixel values representing an image to numerical analysis, usually using a computer. Computer-assisted image analysis procedures include data preprocessing (image sampling and reconstruction, noise removal and reduction, and removal of image blur and other distortions; Billingsley 1983), pattern recognition (Haralick and Fu 1983), the correction of geometric distortions in images (Bernstein 1983), digital filtering for edge enhancement, histogram manipulations for contrast enhancement, and classification of image characteristics through clustering analyses (Estes et al. 1983). Many of these operations have already been performed on Landsat MSS digital data when it is received from EROS. In addition, digital data can be processed using any other numerical or statistical procedure that can be performed on matrix data, and in this manner pixel spectral intensity values can be compared with other values (for instance, observed densities of archaeological discoveries or materials). The archaeological applications of remote sensor data to projection and prediction discussed later in this chapter have used either cluster-based classifications of pixels or raw pixel data and consist of comparisons of these image data with archaeological data distributions.

Some Limitations of Archaeological Remote Sensing

Several archaeological experiments in which remote sensing data and methods have been applied to projection and prediction will be discussed below. The potential utility of remote sensing in such research is high, but a short discussion of some of the limitations of remote sensing in archaeology is necessary for at least two reasons. First, it has been suggested in several places recently that some archaeological remote sensing enthusiasts may have oversold the potential of this body of techniques and methods (Dowman 1980, 1983; Dunnell 1980; Evans 1983a, 1983b; Fuller 1983; Whimster 1983a, 1983b). Second, it is necessary to emphasize that the limitations of any measuring technology are dependent on the conditions under which it is employed, and that the failure of techniques to reach their full potential in one situation does not mean that they will always be less than useful.

Limitations in archaeological remote sensing can be the result of many factors. They may be inherent in the sensing systems themselves; the scale and spatial resolution of data provided by a system impose limits on what can be seen or analyzed. Lenses, shutter speeds, scanning rates, and the speeds and altitudes of the platforms that bear remote sensor devices can impose restrictions on the usefulness of data for specific purposes. Spectral resolution is another important system limitation, and for any purpose it is important to determine just what portions of the electromagnetic spectrum should be measured before remote sensor data are collected. Photographic sensors image only a small portion of the electromagnetic spectrum, but they possess much higher resolution than most multispectral scanner systems.

Instruments available for laboratory analysis may impose another set of limitations on the application of remote sensor data to archaeological problems. While acceptable pocket stereoscopes can be purchased for about $30, more useful stereoscopes can cost thousands of dollars and may not be available to all archaeologists. Digital image-processing systems are even more expensive, although it may be possible to rent time on such systems. Several examples of less-than-optimum digital image analysis being applied to "predictive modeling" in an attempt to save money will be summarized later in this chapter. In many if not most cases, archaeologists who wish to incorporate remote sensing methods into their projects will do better to contact qualified and well-equipped archaeological and cultural resource remote sensing consultants, rather than to entertain notions of doing their remote sensing work "in house."

Environmental factors impose another sort of limitation on archaeological remote sensing. Clouds, mist, and haze can obscure the view of most sensor systems; heavy snow or vegetation cover may also defeat some systems (multispectral scanners and photographic sensors) but have little effect on others. (Radar, for instance, penetrates vegetation canopies with relative ease.) The phenomena recorded by some scanner systems, in particular thermal scanners, are transient and often can be detected only for a few hours or even minutes when conditions are optimal; identifying such optimal conditions may take years of experimentation in any study area (Perisset and Tabbagh 1981; Tabbagh 1977).

Individual human limitations, such as ability or inability to perceive stereo images, experience in photointerpretation, previous familiarity with a study area, or knowledge about what various characteristics of the environment look like, can affect the application of remote sensing to any problem area. In general, in order for a researcher to apply remote sensing to a problem successfully the problem must be stated explicitly; the place of remote sensing in the solution of that problem must be defined; and appropriate methods for discovering, collecting, and analyzing the archaeological (dependent) data that are to be contrasted with environmental (independent) variables must be selected. As we will see when we examine the ways in which archaeologists have applied remote sensing to the problem of predicting the locations and characteristics of archaeological materials, failure to meet these conditions may be one of the most obvious reasons for the lack of satisfying conclusions. It may, in fact, explain much of the present lack of success in predictive modeling in general. Again it should be emphasized that the specialized technology of remote sensing—and the problems it can or cannot help the archaeologist to solve—are best assessed and implemented through a team approach incorporating not only in-house cultural resource management and archaeological personnel, but a specialist in archaeological and cultural resources remote sensing as well.

CONTEMPORARY APPLICATIONS OF REMOTE SENSING TO ARCHAEOLOGICAL PROJECTION AND PREDICTION

A Taxonomy of Predictive Archaeological Remote Sensing

In a previous publication (Ebert 1984:341) I proposed a taxonomy that distinguishes between archaeological sampling, projection, and prediction. But taxonomies are problem-specific, and the problem that I was addressing in this previous publication was the application of remote sensing to survey archaeology as a whole. The purpose of this chapter is somewhat different: it deals specifically with remote sensing applications to projection and/or prediction. As is evident in Chapter 4 of this book, I think it is probably most productive to view *prediction*, here, as an integral part of the explanatory framework of archaeology (see Figure 4.1), as something that archaeologists must do to draw testable expectations from *models* that describe the ways in which we think the archaeological record is related to the organization of past human systems. The term *projection* has been used in the taxonomy in Chapter 4 to designate empirical generalizations about the occurrence of archaeological materials in unsurveyed or unsampled areas on the basis of known distributions in surveyed areas. Because lax definitions can lead to problems in any scientific endeavor (see the Chapter 4 discussion of the *site* concept, for example), the definitions of projection and prediction set forth in Chapter 4 will be used here, rather than those I proposed earlier (in Ebert 1984).

Another theme of Chapter 4 is that we almost certainly do not know how to do successful predictive modeling, in the sense of being able to make generally applicable statements about the location of archaeological materials in unsurveyed areas, at the present time. What is more, we do not know exactly why we cannot successfully predict in a general way. For this reason, it is my belief that almost all of the archaeological "predictions" that have been attempted, by archaeologists in general and by those employing remote sensor-derived data, are probably projections in the sense that this term is used in Chapter 4.

The following discussion of approaches that have used remote sensing data to generate projections is arranged according to a taxonomy that emphasizes differences in *(a)* the things that archaeologists want to predict and *(b)* the remote sensing analysis method employed.

The first taxonomic category comprises approaches that generalize from extant archaeological and environmental data about areas in which archaeological materials are likely to be found but consider only peripherally where materials will not be found. Such approaches could be thought of as prospecting, and their goal is to streamline the discovery of archaeological materials in order that those materials may be studied or preserved. The two basic analytical methods that have been used by archaeologists engaged in this sort of projection are visual analysis and digital analysis.

The second major taxonomic category consists of approaches to archaeological projection that use remote sensing to identify areas where archaeological materials can be expected and areas where they are not expected to be found. In effect, these approaches lead to projections of the differential densities of archaeological materials in a study area or, in some cases, densities of specific rypes of materials. They can also be used to design sampling stratifications that are intended to provide this type of density information. Again, a distinction will be made between approaches that use visual analysis and those that use digital analysis.

What follows is a review of the literature concerning archaeological projective attempts incorporating remote sensing data, organized by these taxonomic categories. The successes and failures of these approaches will be discussed once the summaries have been presented.

First, it should be pointed out that the distinction between these two methods is really technological—people are involved in making decisions whether the processing is done by the human brain or partly by a machine. There are, however, some basic quantitative differences between visual interpretation of environmental variables and digital analysis. One of the most obvious of these is that people generalize when they interpret things from remote sensor data, such as aerial photographs or Landsat visual images. A large, relatively homogeneous area of (for instance) pine forest is identified as such by a human interpreter, and tiny inclusions of oak are ignored. In the course of a computer digital analysis, on the other hand, each pixel is classified, and if an oak pixel falls within a mass of piñon pixels, it is classified as oak forest. In many cases, there is nothing wrong with or unworkable

about the generalizations of human interpreters; the presence of a few oak trees within the piñon forest probably constitutes environmental variation at a scale incomparable with the scale of human mobility and systems organization. For other purposes—in computing an environmental diversity index using a moving filter across space, for instance—the ungeneralized, digital classification may be the only workable data representation.

It has been asserted (Baker and Sessions 1979) that digital analysis is superior to human visual interpretation because human biases are not injected into digital products and because digital analyses are replicable. This is a somewhat optimistic interpretation of what digital analysis actually entails. In human interpretation, subjective decisions are made about where boundaries fall, while in digital analysis subjective, human decisions must be made about the limits of *cluster* boundaries (in multidimensional analyses) or about the confidence limits one is willing to accept as representing useful correlations between the occurrence of cultural and environmental variables. The *meaning* assigned to interpreted or digitally derived variables is subjective in both cases.

Nonetheless, a distinction will be made below between those "predictive" attempts using visual interpretation and those using digital analysis. This is done for the most part for historical reasons, as visual interpretations for archaeological purposes were attempted earlier than machine-processing-based attempts. Digital processing can be cost-saving when large geographic areas are being inspected, and digital-format predictive products are also easier to incorporate into geographic information systems. For these reasons, digital-format products are likely to be the major thrust of remote-sensing-aided archaeology in the future.

Archaeological Projection Through Visual Analysis of Remote Sensor Data

Archaeologists have been using remote sensing, particularly aerial photointerpretation, for the discovery and inspection of sites since the early 1900s (Beazeley 1919; Capper 1907; Lindbergh 1929). Especially in Great Britain and Europe, most archaeological uses of aerial photographs are still directed toward actually seeing the manifestations of sites and structures through shadow or crop marks (Riley 1980, 1982; Wilson 1982). The examples of "prediction" of areas likely to contain positive archaeological evidence discussed here, however, are somewhat more indirect. In these examples the experimenters seek not to see actual sites but rather to correlate the distribution or occuttence of known archaeological materials with certain landform and environmental characteristics. These independent variables are then sought in areas thar have not been archaeologically investigated, and uninvestigated areas exhibiting such properties are postulated to have a high likelihood of containing archaeological materials. In these studies, remote sensing typically provides the basis for characterizing the environment in areas known to contain sites as well as for finding unsurveyed areas with the same characteristics.

One of the earliest predictive studies of this type was carried out in Iowa and used aerial photographs to define and map soil types that were thought to be ideal for the agricultural subsistence practices of the mound-building people who occupied the area shortly prior to European contact (Tandarich 1975). Soil types were photointerpreted and classified according to Department of Agriculture criteria, and those soil types that had been found to be associated with mound sites in the past were further interpreted in stereo to find mounds.

Another early predictive/prospecting study made use of Landsat (ERTS-I) visual multispectral scanner (MSS) data. In this study, Cook and Stringer (1975) attempted first to see the actual manifestations of large, known, historically abandoned village sites in a boreal forest around Kaltag, Alaska, but they were unable to determine whether the spectral signatures they saw indicated the villages themselves. Then, by characterizing the landscape and vegetation in the vicinity of the known village sites, they attempted to predict the *potential* presence of additional villages in other parts of their 450 mi^2 study area. They felt that they were able to relocate 5 of the 12 known village sites, and they also predicted a number of other potential sites, although these were not field checked.

A similar though more rigorous method was adopted by archaeologists at the National Park Service's Remote Sensing Division in a study directed toward locating areas within Shenandoah National Park in Virginia, which had a high potential for prehistoric and historical archaeological site occurrence (Ebert and Gutierrez 1979a, 1979b, 1979c). One impetus behind this study was the desire of park personnel to find exemplary archaeological sites that could be used in interpretive programs. Additionally, this experiment was undertaken in an attempt to show that aerial remote sensing could be of value in the eastern deciduous woodland; a persistent theme in critiques of archaeological remote sensing is that it is only useful in the arid Southwest, where sites can be seen because of sparse vegetation cover. In the Shenandoah project it was not sites themselves that were seen, but rather their settings.

The first step in this project was the selection of environmental indicators (Ebert and Gutierrez 1979b:7), which were chosen not because of any assumed preferences on the part of prehistoric and historical occupants of the area but rather because these environmental characteristics could be photointerpreted from 1:12,000 scale color transparency aerial photographs of two areas of the park. Values for the variables of slope type, slope angle, slope aspect, vegetation type, vegetation diversity, soil thickness, type of surface deposit, bedrock type, and proximity to contacts, faults, and shear zones were formulated, and recognition criteria for each value were explicitly identified.

The next step in the project was to mark the exact locations of previously located historical and prehistoric sites on the aerial photographs. In no case could the site itself be seen, but topographic factors allowed map locations to be transferred to the photos accurately. Within an arbitrary radius of 250 ft around each known location, the environmental indicators were interpreted using a Bausch and Lomb roll-film stereoscope with magnification up to 20×. The results of this

interpretation were tabulated and coded, and the characteristics of places where sites were likely to be found were summarized. Some successful "indicators" of sites were slope angles from 0 to 5°, exposures of southeast to southwest, and proximity to fault or shear zones; historical sites differed from prehistoric ones in that most known historical sites occurred on locally flat areas on sideslopes and in colluvial or alluvial deposits.

Finally, the aerial photographs were reinterpreted to locate areas that exhibited these site-likely indicators but had not been surveyed for archaeological materials in the past. A field check at Shenandoah revealed the presence of previously unrecorded archaeological materials, some of a spectacular nature (including a large nineteenth-century mill site; Figure 9.1), in 45 percent of the projected "likely" areas. One obvious weakness of this study was that no unlikely areas were field checked to test the rejection potential of this projection.

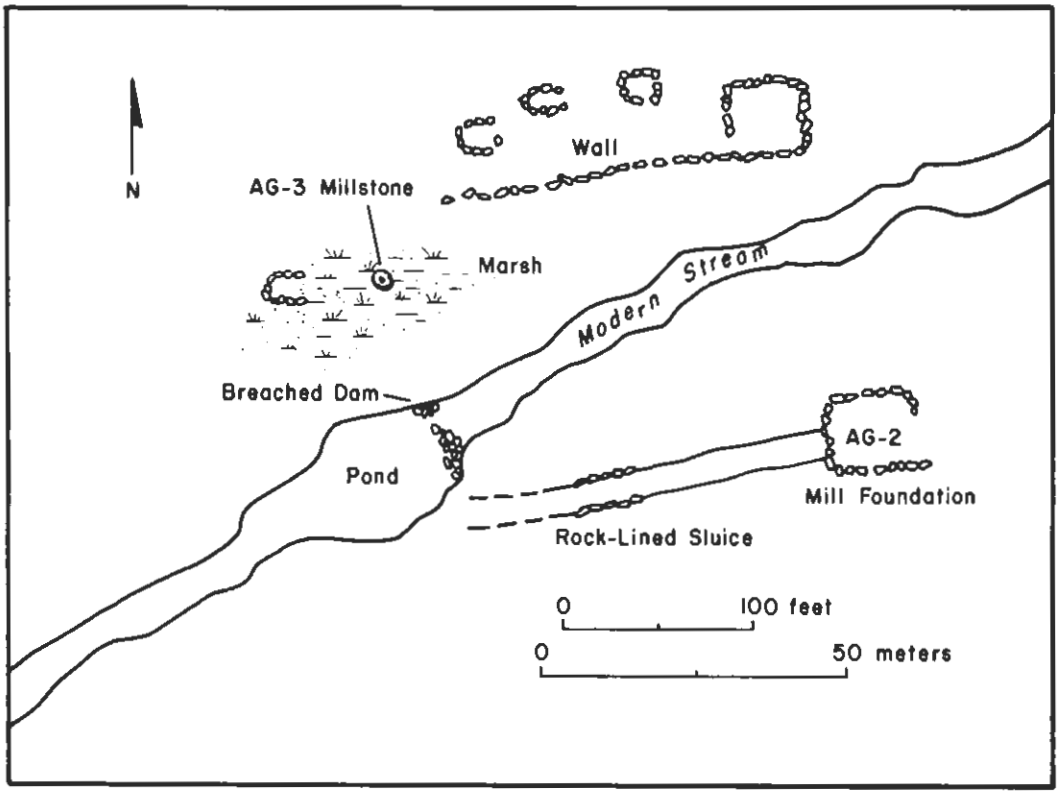


Figure 9.1. An eighteenth-century mill or industrial site discovered in Shenandoah National Park, Virginia, through remote-sensing-aided archaeological projection. The existence of this complex was known from tax records, but its location was not pinpointed until field checking of "probable site areas" derived through the analysis of 1:12,000 scale color infrared aerial photographs of portions of the park was initiated (after Ebert and Gutierrez 1981).

In another projective study carried out during the same year, 1:60,000 and 1:120,000 scale color infrared transparency photographs were inspected for indicators of site occurrence in the National Petroleum Reserve-Alaska on Alaska's North Slope (Gal 1979). Gal felt that, while there was little hope of actually seeing village sites themselves, there was more potential for "identifying areas where *not* to look for archaeological sites and areas of high archaeological potential" (1979:1). He sought such indicators as the consistent appearance of whaling lanes and areas where early melting of snow and ice provided locations desirable for springtime camping grounds. Areas with known archaeological sites appeared to be lighter in color than surrounding areas in the color infrared aerial photographs, which were taken in July; Gal believed that this indicated better-drained places where vegetation flourished but died off first. Gal concluded that such studies held great potential, especially in the Arctic where ground reconnaissance is expensive and difficult and where "narrowing down" of survey areas is virtually necessary.

Two studies that followed the lead of the Shenandoah projection experiment were also undertaken in the Eastern forests by archaeologists from the National Park Service's Southeastern Archaeological Center. Inspection of color infrared aerial photographs, which make it relatively easy to recognize the distinction between water and land, provided a preliminary indication of where to conduct archaeological surveys in the Big Cypress Swamp in Florida (J. Ehrenhard 1980). In such areas, of course, human occupation takes place only where there is no standing water, a criterion that restricts "site likely" areas severely. By noting the locations of small mounded areas surrounded by sawgrass and water, archaeologists were able to narrow down their survey efforts to a very small percentage of the total area encompassed by Big Cypress Swamp. A more complex series of indicators interpreted from aerial photographs, including topographic, hydrologic, and soil productivity variables, were correlated with different temporal and functional characteristics of a sample of previously known archaeological materials in the Chattahoochee River Recreation Area; the resultant model proved to be successful in locating sites from different time periods (E. Ehrenhard 1980).

Digital Approaches to Archaeological Projection

A digital approach to detecting and analyzing the "residual effects of prehistoric human settlement upon landscapes" was undertaken in the late 1970s in southwestern New Mexico in an attempt to locate Animas phase pueblos for further study (Findlow and Confeld 1980:31). Landsat MSS computer compatible tapes (CCTs) were analyzed at Columbia University using Map 1 software. The spectral characteristics of "catchments" of 32 by 40 pixels (about 1200 acres), 16 by 20 pixels (about 300 acres), and 8 by 10 pixels (80 acres) centered on 8 large (100–500 room) Animas phase sites and 33 randomly selected points that had not been previously surveyed were compared using analysis of variance statistics. Findlow and Confeld concluded that soil and vegetation were darker around site areas than in nonsite

catchments (Figure 9.2) and that these differences were particularly obvious in the 8 by 10 pixel examples. The lower reflectance was attributed to greater moisture retention and to the existence of cultural debris in middens surrounding the large sites.

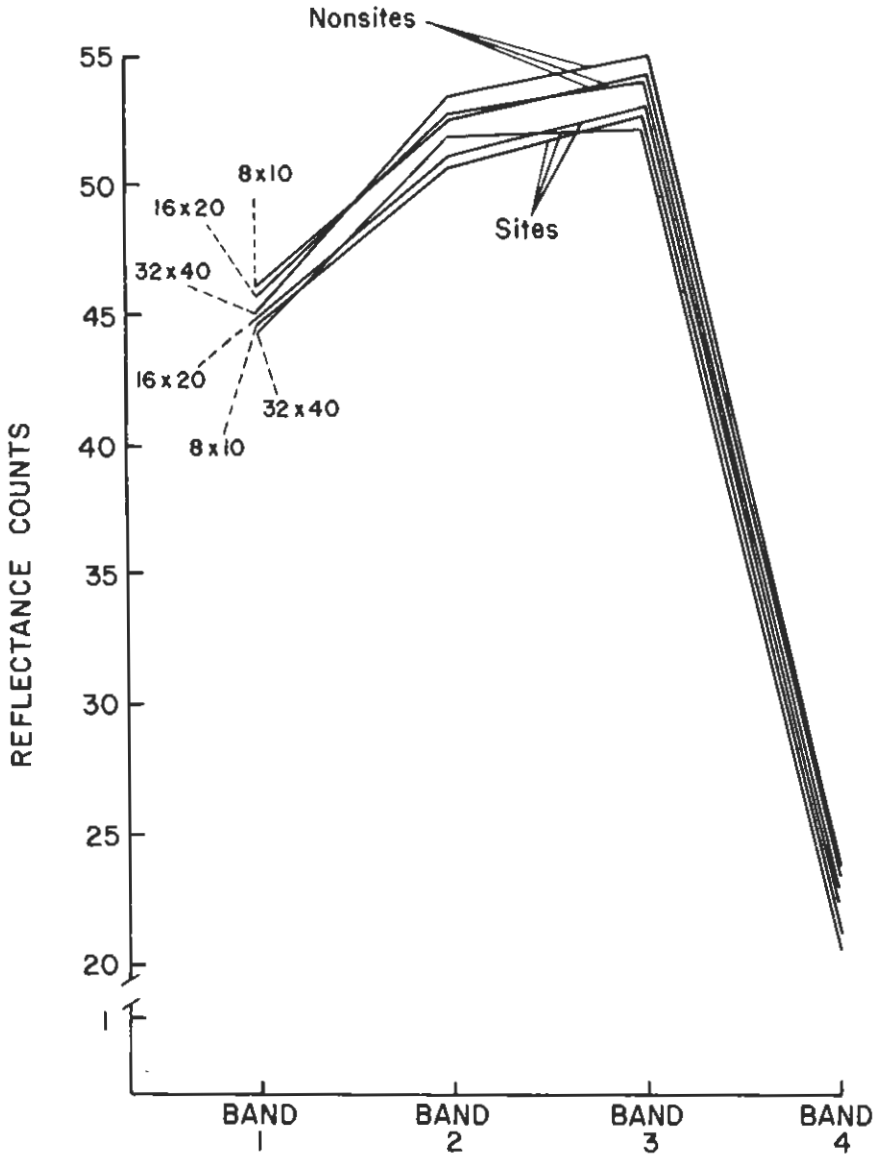


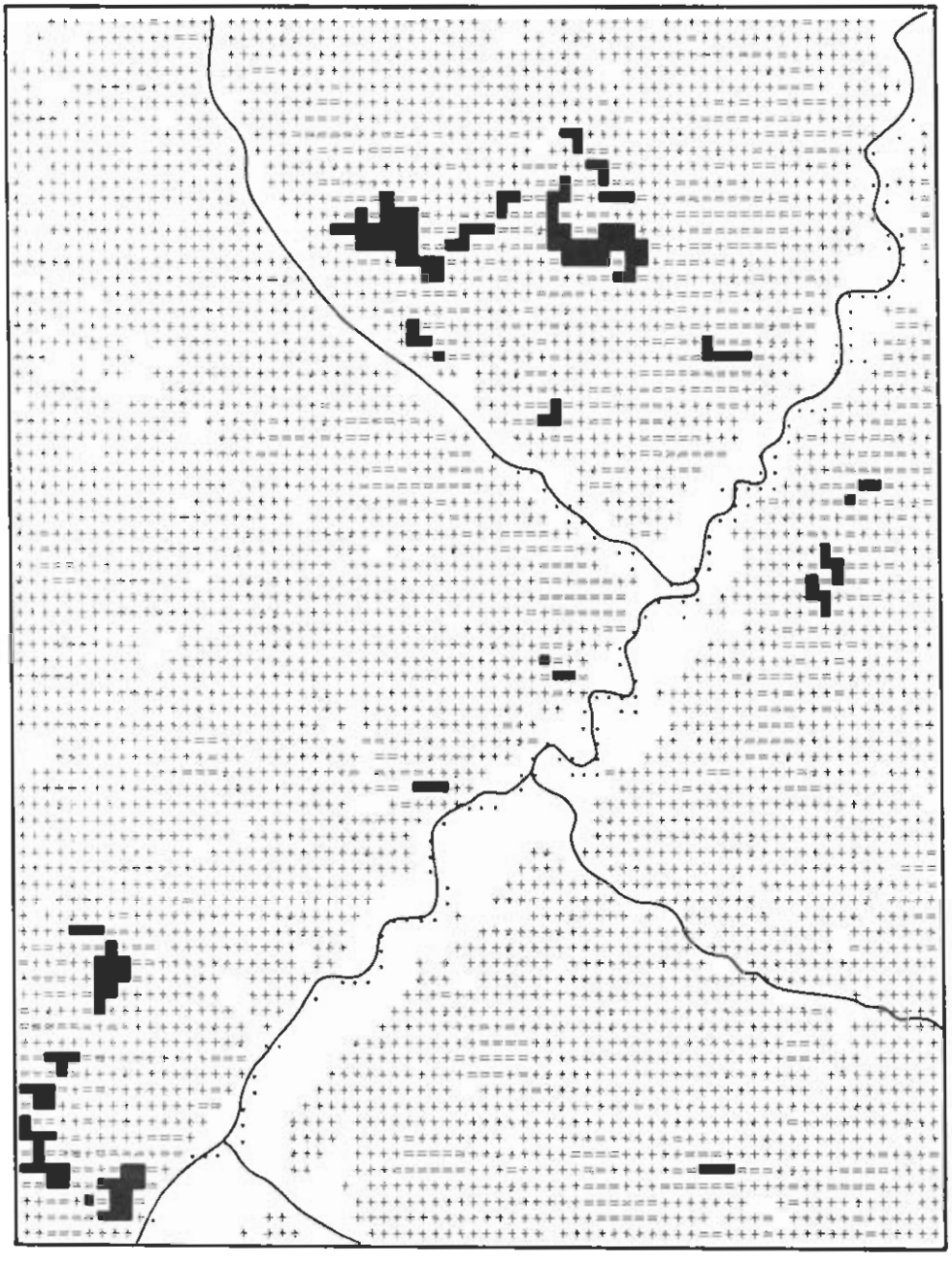
Figure 9.2. Digitally derived reflectance values for sites vs nonsites in southwestern New Mexico taken from Landsat MSS computer compatible tapes (after Findlow and Confeld 1980).

It is difficult to determine whether the above example actually constitutes a case of "seeing sites," or whether it was the contexts of the sites that were being detected. Another projective study utilizing remote sensing that evokes the same question has been pursued since the early 1980s by K.-Peter Lade at Salisbury State University in Maryland (Lade 1981a, 1981b, 1982). Using Salisbury State's ASTEP II software for the analysis of Landsat MSS data, Lade examined the land cover and geology of a 34,000 km² study area (an entire Landsat scene), classifying pixels on the basis of "angular distance relationships observed in vector space of normalized data" (principal components analysis; Lade 1981a:13). He found that dry, sandy ridges were easily discerned through density slicing of Landsat's band 7 (Figure 9.3), and that such ridges were usually entirely covered with cultural materials representing multiple occupations through long time periods. It is not clear whether Lade is identifying the effects of such occupation or a particular landform type conducive to occupation (or to finding materials—sand ridges typically have discontinuous vegetation cover), but his projections are undeniably successful at finding site-likely areas.

A more rigorous prospecting approach, which was also carried out in an eastern coastal plain and piedmont setting, is Wells's (1981) study, which is explicitly based on discrimination of landscape features. Wells selected "predictive environmental variables" (1981:22), including distance to water, specific soil types, and specific geomorphological and topographic settings, as well as known archaeological site locations, and subjected these variables to a logistic regression. His results were tested by field-inspecting both site-likely and site-unlikely areas. Although Wells primarily used information derived from map-based geographic information systems, based on photointerpretation by others, he discusses at length the potential for automatic projections of this type using Landsat MSS data.

An approach similar to the earlier Shenandoah experiments was used in Kentucky by Carstens et al. (1981). Stereo photointerpretation of 1:20,000 aerial photographs was performed by a number of independent interpreters, and the results consisted of codings of the characteristics of landforms and vegetation cover in a 400 by 400 m grid overlain on the photos. The same exercise was then undertaken using 1:7920 scale photographs and a 100 by 100 m grid overlay. The smaller grid overlay interpretations proved to be more useful for identifying areas in which archaeological sites were found (using a previously known sample that presumably had *not* been inspected by the interpreters prior to their inspection of the photographs), resulting in the recognition of 13 of 19 known sites (68 percent accuracy). A field check revealed that additional, previously unknown sites could actually be found in 78 percent of the projected likely grid cells. Another study following almost the same methodology but using photomosaic (monoscopic) interpretation rather than stereo photointerpretation (Haase 1981) predicted site densities on Cedar Mesa, Utah, with more variable results.

An ongoing projective experiment using known locations of Gallo-Roman villas in the Burgundy region of France (Madry 1983, 1984) is especially interesting in that it incorporates modern digital analysis in an area that had until his study



SAND RIDGE LOCATIONS--IR BAND 7

Figure 9.3. Sand ridge locations in Maryland with high probabilities of archaeological site occurrence. Landsat MSS band 4 (near infrared) data were digitally analyzed to arrive at this map of areas in which sites are likely to occur—or, more precisely, are likely to be found by archaeologists (after Lade 1981a).

been a stronghold of the "seeing sites" approach to remote sensing; usually aerial photography from light aircraft served only as a peripheral means of documentation. Madry analyzed two Landsat 2 MSS scenes using the General Electric "Image 100" system, with known villa locations serving as a training set (that is, the spectral characteristics of known villa locations served as "instructions" to the computer, which then selected other areas with similar spectral signatures). He has concluded thus far that the resolution of the Landsat data (80 by 80 m pixels) is too gross to allow new villa sites to be found, although relatively intense and continuous land use since the Gallo-Roman occupations may also be a factor in his lack of success.

All of the approaches summarized above share a number of common characteristics. Their primary goal is the identification of areas that are likely to contain archaeological materials, based on characteristics of the locations of previously discovered sites. While they are strictly empirical, these studies are also "positive" in that their goal is to find sites or materials for archaeological study. Their results cannot, therefore, be easily converted into statements about areas where sites and materials *will not* be found and thus areas that can be ignored for clearance or study purposes.

The more explicitly "predictive" studies that will be discussed below are, for the most part, extensions of these projections. Although such extensions are useful, the resultant models are no more explanatory than the correlations on which prospecting for site-likely areas are based.

"Predictions" of Site Occurrence/Nonoccurrence or Site Densities Based on Remote Sensor Data

Unlike the prospecting approaches to predicting likely areas in which to find undiscovered archaeological materials discussed above, the avowedly "predictive" remote sensing approaches to the archaeological record that are summarized below are directed toward identifying areas of differential occurrence of sites within regions. Such differences may be expressed in terms of sites vs no sites (or nonsites), differential densities of sites, or variation in densities of more than one temporal, cultural, or functional site type between zones within a study or management area. Nonetheless, the discovery of these differences is approached in essentially the same way as was site "prospecting" in the section above. The locations of known sites, or of different types of sites, are tabulated from previous survey data; the study area is then divided through either an arbitrary stratification (e.g., grid cells) or an informed stratification (environmental zones of one sort or another). Through one of a number of statistical techniques, the physical locations where sites previously have been found are correlated with the physical locations of independent environmental variables (see Chapters 5–8 of this volume). On this empirical basis, projections are made about where sites will or will not be found and about the densities of sites in general or of different types of sites in areas where the archaeological record is not known.

Clearly, these are not predictions in the explanatory sense set forth in Chapter 4; they are inductive, empirical generalizations or projections. Some of the implications of this fact for the utility of such predictions are discussed in a subsequent section of this chapter, along with ways of going beyond correlations of archaeological manifestations and environmental variables.

Archaeological Prediction and Visual Interpretation

To my knowledge, the first "predictive" archaeological attempt utilizing remote sensing as a major environmental data source was initiated in early 1977 as part of the National Park Service Alaska Area Office's cultural resources assessment of the National Petroleum Reserve in Alaska (NPRA). This experiment was originally envisioned simply as an exercise in sample design; it has obvious implications for remote-sensing-based archaeological prediction, however, and in fact the methods used were incorporated in almost identical form into an avowedly "predictive model" of site densities in the San Juan Basin of New Mexico that was carried out by the National Park Service shortly after the Alaskan project was completed.

The National Park Service was asked to conduct a reconnaissance of the NPRA, which covers about 23 million acres (92,000 km^2) of Alaska's North Slope, by the Bureau of Land Management prior to the opening of the area to virtually unrestricted petroleum exploration. The ideal would have been to survey a representative sample of the whole project area, but this was nearly impossible given the short, 8–12 week summer field survey season; the general inaccessibility of the study area; the impossibility of land transportation during times when the ground was not snow-covered; and the tremendous area to be covered. Although the North Slope appears almost featureless on atlas maps, it extends from the peaks of the Brooks Range across foothills and a sand-mantled upland region to the poorly drained coastal plains. The great environmental variability and logistical problems of surveys in the Arctic meant that any sort of successful appraisal of the nature and distribution of archaeological materials in the NPRA would require careful sampling and planning, and the National Park Service's Remote Sensing Division in Albuquerque was asked to provide assistance.

The most interesting potential use of remote sensor data for this project was as a basis for sample stratification and design. A sample design was created, based on an informed stratification of the NPRA into a relatively small number of "ecologic-cover type zones" (Brown and Ebert 1978; Ebert 1978; Lyons and Ebert 1978) determined on the basis of visual interpretation of Landsat MSS color composite images.

An initial ecologic/cover-type stratification was compiled through visual interpretation of 10 Landsat scenes and mapped on a 1:500,000 scale base map of the NPRA (Figure 9.4). Subsequently, 1:60,000 and 1:120,000 scale color infrared aerial transparencies, which present a spectral picture nearly identical to that of Landsat color composite MSS visual images, were used as a preliminary check on the

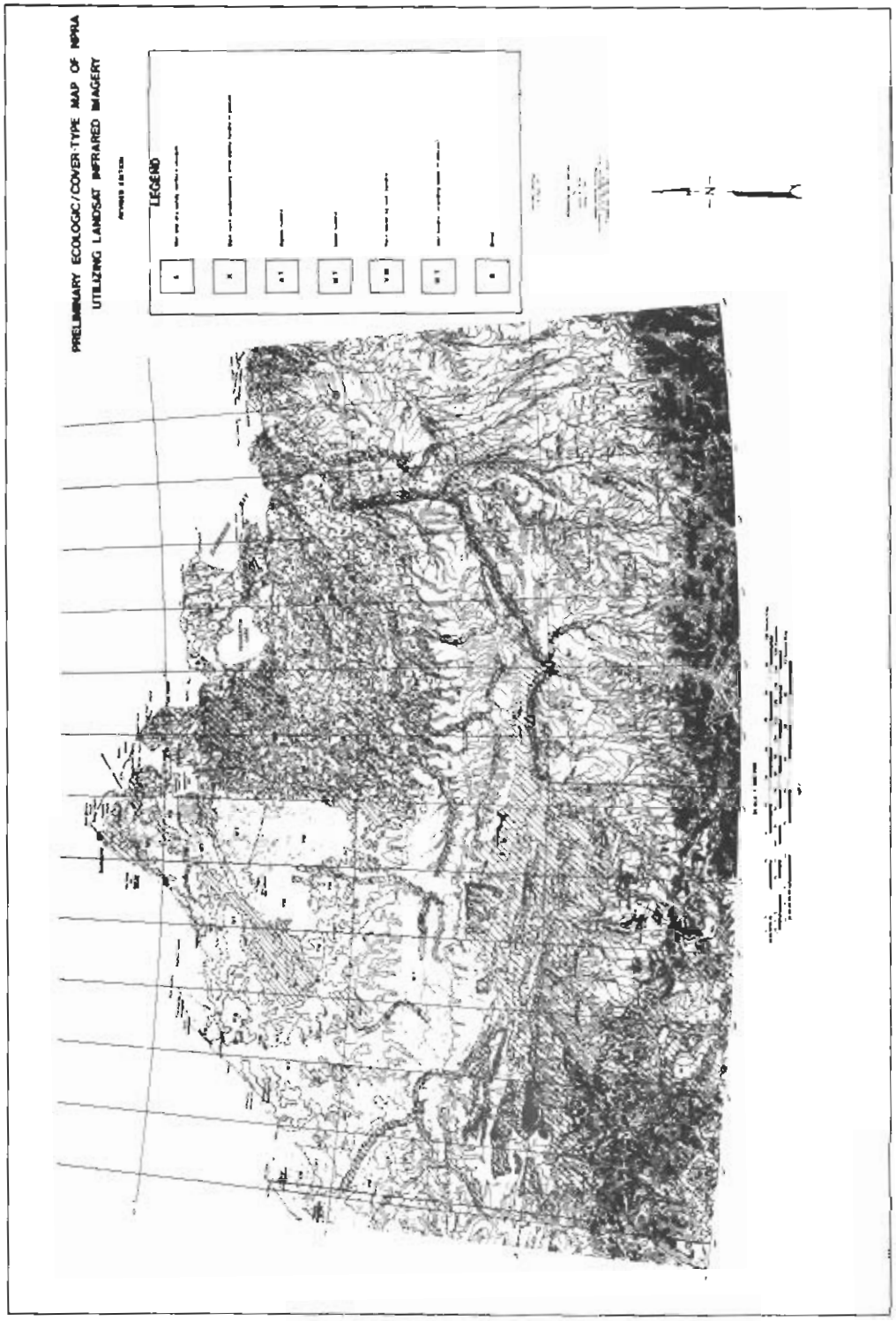


Figure 9.4. An ecologic/cover-type map of the National Petroleum Reserve in Alaska (NPR-A) compiled with the aid of interpretation of Landsat color composite visual data. The seven cover types, which are composites of hydrologic, vegetative, and topographic indicators, proved to be successful in the discrimination of differential archaeological site densities and functional/temporal site types (Ebert 1978:Fig. 2).

Landsat-derived classification, which defined seven zones based primarily on interpreted drainage and vegetation-cover differences. Although the small-scale aerial photographs could also have been used to stratify the NPRA for sampling, this would have required the interpretation of literally tens of thousands of frames, a practical impossibility. The second level of ground truthing consisted of an examination of zone boundaries across the NPRA from a helicopter platform, which illustrates the point that ground truthing does not always have to be done on the ground. Additional oblique aerial photographs were taken during this helicopter examination, using 35 mm cameras and color infrared film for documentation purposes (Ebert 1980). Finally, observations of boundaries and the vegetative composition of each ecologic/cover-type zone were made on the ground.

Although it had been assumed that this stratification would be used to select those areas in which survey would be carried out by NPS field crews, conflicting ideas about the goals of the survey prevented this from being accomplished. By the time the preliminary stratification had been completed, the NPRA survey crews were already in the field and had already selected areas to be surveyed based on potential site densities—that is, areas that were believed, on the basis of past experience in the Arctic, to be likely to contain concentrations of archaeological sites were chosen for reconnaissance. As pointed out previously, this is a valid approach if it is the highest concentrations of spectacular sites that one is seeking, and in fact the major outcome of the NPRA cultural resources assessment was the setting aside of a number of National Register districts with high concentrations of archaeological materials.

Even though the remote sensing sample stratification was not used to select survey areas, the discriminatory power of the sample stratification was tested using the data that were collected. The approximate boundaries of the surveyed areas were marked by the survey crew leaders on 1:250,000 topographic sheets, and the survey areas were then carefully stratified, using detailed versions of the ecologic/cover-type zones discussed above. The area of each stratum actually surveyed was measured with a digital planimeter and compared with the numbers of the types of sites discovered during that survey. On the basis of this information, purely empirical "predictions" of site density within particular strata were made.

The second season of survey, carried out in the summer of 1978, was also conducted without reference to the ecologic/cover-type sample stratification. Portions of four strata that were partially covered during the first season (summer of 1977) were also surveyed during this second reconnaissance. A comparison of site densities in these strata between the two field seasons is interesting (Table 9.1). The striking differences may be a result of variations from place to place within the NPRA in the effectiveness of the ecologic/cover-type stratification. Alternatively, these differences may reflect changes in the ways things were sought in the field, in the experience and expectations of the crew in successive summers, and in the ways that sites were recorded. Moist tundra, where the lowest densities were found in both seasons, is typically covered by dense grass tussocks, and none but the most obtrusive archaeological materials can be found there. The "brush" stratum occurs

TABLE 9.1.

Sites located by the NPR-A cultural resource surveys, 1977 and 1978

Stratum	*Area Surveyed (km)*	*Number of Sites*	*Sites/km²*
	SUMMER 1977		
Moist tundra	428.85	34	0.079
Alpine tundra/moist tundra	536.83	126	0.253
Brush/moist tundra	117.73	29	0.246
Bare rock/gravel	94.66	76	0.803
	SUMMER 1978		
Moist tundra	56.48	12	0.212
Alpine tundra/moist tundra	219.82	110	0.500
Brush/moist tundra	76.91	73	0.949
Bare rock/gravel	46.42	18	0.389

along rivers and lakeshores, and it is in this zone that large village sites are usually found, probably because of the availability of firewood. Sites do not occur everywhere within brush areas, however. The brush cover must occur in conjunction with one or more of a number of geographic situations (caribou crossings, river confluences, the windward side of lakes, etc.) if the likelihood of finding a site is to be increased. The survey crew may well have learned to identify these combinations of factors, which would account for the dramatic increase in identified "brush" sites during the second season. Another possibility is that, while the "brush" strata in which survey was carried out in the first and second survey seasons were of the same composition, other properties of the strata, such as distances to boundaries or sizes of portions of this stratum, may have been different (Michael Garratt, personal communication 1985). This might underline the appropriateness of attempting to derive diversity or heterogeneity measures from remote sensor data for predictions, a topic that will be discussed at length later in this chapter.

At about the same time that the National Petroleum Reserve in Alaska cultural resource project was winding down, the National Park Service's Southwest Regional Office became involved in studying cultural resources as part of another multi-agency impact assessment, the San Juan Basin Regional Uranium Study in northwestern New Mexico. The Bureau of Indian Affairs, which administered the study, requested that the NPS Southwest Regional Office study the potential impacts of uranium mining and associated development on the cultural resources of this 100 by 100 mi area.

The primary task undertaken by the National Park Service for this purpose was the consolidation, in consistent format, of all available archaeological survey data from some 4000 known surveys that had taken place in the San Juan Basin, a herculean task in itself. Extensive data on more than 16,000 sites were compiled and recorded on computer media, and software was devised to make access to any aspect of these data simple and economical. These data have formed the basis of a wide

range of assessments and discussions of the archaeology of northwestern New Mexico and the dangers threatening these resources today (see especially Plog and Wait 1982).

As part of this impact assessment, the National Park Service's Remote Sensing Division was asked to attempt to predict distributions of sites—to "make some statement about the distribution of archaeological sites throughout the Basin" (Drager and Lyons 1983:2) using remote sensing data and methods. An approach virtually identical to that of the NPRA remote sensing sampling design was adopted for the San Juan Basin project. An ecologic/cover-type stratification was prepared through the visual interpretation of Landsat MSS color composite visual images, based on the methods used in the NPRA (Camilli 1984; Drager 1980a, 1980b; Drager et al. 1982). Eighteen different cover types with 22 additional subtypes were defined for this area, which is environmentally far more complex than Alaska's North Slope (Figure 9.5). In addition, these cover types were cross-correlated with eight landform types (see Drager and Lyons 1983 for details). The resultant zones were mapped on a 1:250,000 scale base map. Other information also examined for the San Juan Basin included surface geology and average annual precipitation.

The first step in making projections about site densities was to overlay 2 by 2 km grid squares to code the previously surveyed areas onto an ecologic/cover-type map of the basin. Surveyed squares that comprised more than one ecologic/cover-type zone were eliminated. Numbers of archaeological sites found within each zone in the course of previous surveys were then determined by searching the computer data base. For each zone, the total number of sites found was divided by the area surveyed to calculate a density figure. The number of sites in each zone was then predicted. Previous archaeological surveys had only been conducted in 21 of the 40 zones/subzones defined during ecologic/cover-type mapping, and predictions were made only for these zones. Still, some 51,700 sites were predicted to be present in these zones, a sizable (and perhaps unmanageable?) number.

Several other projective experiments in New Mexico, all based on the methodology used in the NPRA and San Juan Basin projects, have been reported in the literature (Camilli 1979a, 1979b; Camilli and Seaman 1979; McAnany and Nelson 1982), and an additional experiment in predicting site densities across ecologic/cover-type, surface geological, and soils zones (all based on remote sensing) has since been carried out by the Remote Sensing Division (Drager and Ireland 1986) as well. All of these approaches exemplify the ways in which an area can be stratified into different and often empirically significant areas or strata for sampling or for empirical projection from known site distributions to the distributions of sites in areas not yet surveyed. All suffer the same deficiencies exhibited by other empirical correlative "predictive" schemes: they are not explanatory, and their success or failure at prediction—even if "tested"—cannot be accounted for.

An additional problem of projective or predictive experiments based on the use of archaeological data from many surveys should be mentioned briefly here. Although many states or regions of this country have well-developed data management or geographic information systems from which great volumes of survey data

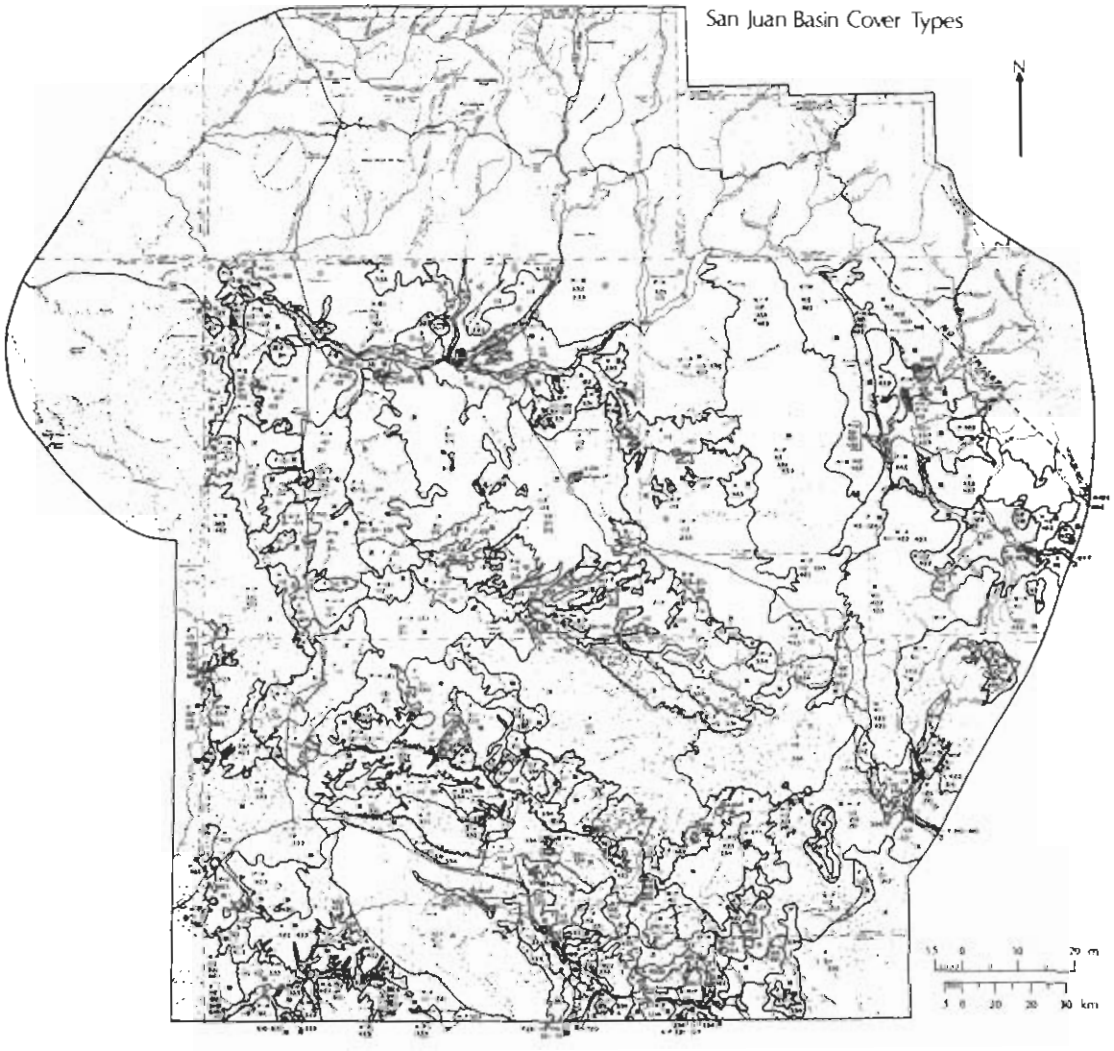


Figure 9.5. San Juan Basic ecologic/cover-type zones delineated through the interpretation of Landsat MSS visual data produced in an attempt to project archaeological site densities and the differential locations of archaeological site types in northwestern New Mexico (Camilli 1984:Fig. 4). The interpretation methods followed in this effort were essentially the same as those used in the NPR-A interpretation shown in Figure 9.4.

Figure 9.5. Continued

KEY

Vegetation

Shrubland

100 Big Sagebrush series
- 111 Big Sagebrush—galleta grass
- 112 Big Sagebrush—blue grama
- 113 Big Sagebrush—western wheatgrass

120 Salt Desert Series
- 121 Shadscale mixed saltbush—alkali sacaton
- 122 Shadscale mixed saltbush—Indian ricegrass
- 123 Shadscale—grama grass
- 124 Four-wing saltbush—greasewood—alkali sacaton
- 125 Four-wing saltbush—blue grama-sand dropseed

Grassland

210 Rice Grass Series
- 211 Rice grass galleta grass
- 212 Rice grass big sagebrush
- 213 Rice grass/mesa dropseed

220 Grama Grass Series
- 211 Blue grama galleta grass
- 222 Black grama galleta grass

230 Alkali Sacaton

240 Meadows

Woodland

310 Riparian woodland (cottonwood, willow, tamarisk)

320 Mountain shrub woodland
- 321 Gambel oak mixed shrubland

330 Pinyon — Juniper woodland
- 331 Pinyon—Juniper
- 332 Pinyon—Juniper Gambel Oak
- 333 Pinyon—Juniper Big sagebrush
- 334 Pinyon—Juniper Blue Grama

Coniferous Forest

410 Spruce-fir forest
- 411 Englemann Spruce Alpine fir
- 412 Douglas fir Colorado blue spruce

420 Ponderosa Pine Forest
- 421 Ponderosa Pine
- 422 Ponderosa Pine Gambel Oak
- 423 Ponderosa Pine Pinyon-Juniper

Deciduous Forest

510 Quaking Aspen Forest

Agricultural Crops and Urban Centres

610 Irrigated Crops

620 Dry Land Crops

630 Urban

Barren and Disturbed Lands

710 Annual Weed Barrens (communities 124 and 125)

720 Malpais

800 Standing Water

Landforms

P Plains — nearly level surfaces with some undulations or dissections

S Sand Dune — a layer of windblown sand in a typical dune pattern

B Badlands — high density of gullies, ravines and sharp-backed ridges; usually clay or shale

A Mesas — cliff bounded areas, higher than the surrounding terrain, with level to gently sloping often dissected surfaces

U Plateaus — generally higher than the surrounding area with level to moderately sloping surfaces with intervening steep to vertical escarpments

H Hill — moderately steep to steep slopes and somewhat rounded relief of less than 500 feet

M Mountain — dominated by local relief in excess of 500 feet: steep slopes and rugged summits

E Erosional area — nearly devoid of vegetation with alkaline or saline soils

can be recovered, all archaeologists are aware that survey data are often inconsistent. The goals of archaeology have changed in a broad sense through time, and many sorts of information recorded today were ignored in the past. Survey data seldom contain any useful indication of survey intensity, a factor that is all-important for judging the completeness of recovery and representativeness of data collection (see discussion in Chapter 4). As discussed earlier, geomorphological and climatic conditions may be as important in determining what is found during surface surveys as what is actually there; most survey data do not provide information on these factors, either. In predictive experiments that utilize data from many different surveys, it may primarily be variations in survey quality, rather than the characteristics of the actual archaeological record, that are being measured. Some ideas about how archaeologists might deal with this problem are presented in Chapter 7.

Archaeological Prediction Through Digital Analysis

A final class of "predictive" experiments utilizing remote sensor data makes use of the computer analysis of digital remote sensor data—either digitally recorded Landsat or other satellite data or, in a few cases, analog (photographic) images converted to digital form. Digital analyses can take a number of forms, distin-

guished on the basis of the degree of explanatory meaning imparted to the results of the analysis. Some of the digital analyses presented here purposely avoid making statements about *why* archaeological sites or materials are found in specific areas or in conjunction with certain spectral value ranges, citing instead the "objectivity" of automatic digital processing. As discussed above, this assessment is not completely realistic, for in any kind of computer processing of image data, decisions about cutoff points for clustering or correlative analyses must be made subjectively.

In order to make this point clear, a typical digital image analysis procedure—clustering analysis—will be summarized briefly here. The goal of digital clustering analysis is recognition of recurrent spectral patterns across multiple bands of a multispectral image (such as a Landsat MSS scene or subscene). Different sorts of phenomena on the earth's surface—plants, water, bare soil, or rocks, for instance—reflect electromagnetic radiation differentially across multiple spectral bands. For example, water absorbs almost all infrared radiation and appears black in Landsat infrared bands and lighter in the red and green bands; bare soil reflects highly in all four bands; and growing vegetation reflects infrared radiation but absorbs light in the red band. Digital clustering analyses examine the differential values of each pixel in more than one band and group them into clusters on the basis of subjectively determined cutoff values.

A digital analysis can be either supervised or unsupervised. In a supervised classification, a human operator directs the computer analysis by specifying a "training set" of areas that represent each desired cover type class to be discriminated. The computer then attemprs to fit the spectral variability within the data into these clusters (not always successfully). In an unsupervised classification, the computer discriminates clusters only on the basis of arbitrary cutoff values that draw boundaries between clusters of values in n-dimensional space (where n is the number of spectral bands used in the analysis). There are several kinds of cluster cutoff boundaries that can be used, including minimum distance to means classifiers (which measures between-cluster centroids), parallelepiped classifiers (which consider the range of variance in a training set), and maximum likelihood classifiers (which evaluate both the variance within classes and the correlation between them; Lillesand and Kiefer 1979:457–487). The machine then tells the operator how many classes it has bounded, and the operator must decide what is actually being represented by each class. Following unsupervised classification, classes are usually collapsed into fewer classes by the operator, and these aggregate classes are named according to what the operator thinks they represent. Subjective decisions obviously enter into each rype of cluster analysis, and the interpretation of what the results of such an analysis mean is always subjective as well. The actual composition of each area can be ground-checked and can also be compared with values of dependent variables (archaeological site densities, in most of the examples summarized in this chapter), but the reasons for the correlation between environmental and cultural variables are not obvious.

An example of a remote-sensing-assisted predictive approach based on cluster analysis is a study of the archaeology of the Bisti–Star Lake region in northwestern

New Mexico, which was done in anticipation of large-scale coal mining (Kemrer 1982). The goal of this project was "to assess archaeological variability on lands [designated] for potential competitive lease coal development" (Kemrer 1982:2). This assessment was based on a sample of the total area and involved the implicit construction of a "predictive model" (Kemrer 1982:2). Archaeological distribution data were derived from six previous surveys that had been undertaken in the immediate project area.

Remote sensing was used to generate independent environmental data against which to compare the Bisti–Star Lake archaeological sample. Two basic assumptions of this project were that "site locational patterning is strongly related to the location of critical environmental resources [and that] it is likely that site frequencies and environmental resources are directly related" (Baker and Sessions 1982:63). The critical environmental variables that Baker and Sessions decided to measure were soil associations and the presence of washes, which they concluded had not changed appreciably since prehistoric times and which directly affect many other variables that might have changed, such as vegetation and the distribution of animal resources. Landsat MSS data and digital analysis methods were chosen because of the size of the study area, the replicability of digital numerical methods as compared to visual interpretations, and the ease of statistical comparison of numerical output values with archaeological site densities. An October 1977 Landsat MSS scene was chosen, and Soil Conservation Service soil mapping units, superimposed on aerial photographs, were used as training samples in a discriminant function analysis performed at the University of New Mexico's Technology Applications Center.

The discriminant analysis, using a maximum likelihood classifier, distinguished eight soil classes, which "was considered adequate for predictive modeling purposes" (Baker and Sessions 1982:66). Based on methods developed during a previous predictive study in New Mexico (Baker and Sessions 1979), a 2 by 2 km grid was imposed on the study area, and archaeological and independent environmental variables were compared within the cells of this grid. Archaeological site density was correlated with four different, and perhaps overly complex, sets of remote-sensing-derived variables, which they describe as follows:

> 1. The eight variables (seven soils associations, plus the category "washes") output by the digital image analysis;
>
> 2. A second set of eight variables based on the proportion of pixels per grid unit classified into each class;
>
> 3. A set of 28 variables that represent all unique two-way interactions between the environmental classes (classes 12 through 78) with values derived by multiplying the number of pixels classified into each member of each two-class set within each grid unit; and
>
> 4. A fourth set also containing 28 variables representing all unique two-way interactions between the eight environmental classes, where the proportional number of pixels classified into each member of each two-class set within each grid unit was multiplied to derive values [Baker and Sessions 1982:68–69].

The stated rationale for developing these four rather complex sets of environmental variables was that it was not known whether cultural variables (i.e., site densities) would vary with absolute or proportional classified pixel frequencies, and that most squares contained more than one environmental class.

The formula used for modeling site-component densities was a linear equation in which observed site densities were taken to be the direct result of summing a set of weighted independent environmental variables (Baker and Sessions 1982:84). Weights and constants were determined through a series of backward stepwise regressions; separate environmental variables were chosen from the 72 variables that best correlated with the occurrence of eight temporal/cultural classes of sites. Regressions were done using squares that contained more than 20 percent total survey coverage, those with more than 40 percent coverage, and those with more than 60 percent coverage. The regression with the squares containing more than 60 percent coverage exhibited the least error, with R^2 values ("explained variance"; Baker and Sessions 1982:87) ranging from 52 to 86 percent for each best-fit variable. This preliminary model, which Baker and Sessions term *Model I*, was used as the basis for making predictions about site-type densities for 813 2 by 2 km grid units, and 15 of these units were then surveyed as a test of the prediction.

Based on these results, another regression model was then generated in an effort to project site densities more accurately. This model showed smaller average error than did Model I, with R^2 values between 52 and 68 percent. Kemrer (1982:98) notes that there are high "correspondences in variables selected between Models I and II," meaning that in general those variables that correlate positively with the occurrence of archaeological sites in one model do so in the other as well, a pattern that holds for a large number of the 72 variables inspected. "Therefore," he concludes, "It is highly likely that the environmental variables are sensitive indicators of site frequency variations."

I would suggest that this correspondence might, instead, be the result of the variables all having been artificially constructed from the original eight remote-sensing-derived soil and wash classes. Such variables cannot be independent, and if patterning exists in the original eight variables then it will also be found in a large number of the 72 derived variables.

Another remote sensing experiment based on the assumption that environmental factors are significant predictors of site locations was conducted in southern Colorado by the University of Utah's Archeological Center in order to assess the prehistoric and historical archaeological materials along a proposed railroad route (Holmer 1982). In this study, "raw" pixel data digitized from a visual Landsat image were correlated with the presence or absence of previously discovered archaeological sites in parts of the study area that had been surveyed, and predictions were then made about the probability of occurrence of sites in those areas not previously surveyed. First, each 128 by 128 pixel portion of a Landsat visual image was digitized or resampled into 500 by 500 m pixels, 22,400 of which were required to cover the entire study area. These pixels were not subjected to a cluster analysis, but rather their spectral characteristics were compared directly with site presence vs absence

through a discriminant analysis. The desired result was not determination of group membership per se, but rather determination of the probability of group membership within the group that contained sites; in this way "sensitivity zones [were] defined by ranges of probability of site presence" (Holmer 1982:37–38). A very small number of "site present" pixels—nine with historical sites and 119 with prehistoric sites—were used to define the dependent cultural variable; eventually the historical site-present cells were dropped, and only the prehistoric cells were used in discriminant analysis. These cells constituted only 0.53 percent of the study area.

Discriminant analysis compared site-presence with three spectral bands and with the ratios between the red and blue Landsat bands. It was found that site presence vs site absence could be best distinguished on the basis of data from the red filter band (Holmer 1982:42). The same data were then compared using logistic regression analysis, and under this procedure the no-filter data (i.e., simple black-and-white density values within each grid cell) were the "best predictor" (Holmer 1982:44). Based on the results of the logistic regression, the total study area was divided into three groups of pixels: those with a greater than 0.275 probability of having sites, those with a probability falling between 0.275 and 0.100, and those with a site probability below 0.100. These three zones were mapped, and the lowest probability zone was classified as the most preferable area for development (Figure 9.6).

Holmer advances a number of conclusions based on this experiment. He suggests that the pixel size used in this study, 500 m by 500 m, was excessive and that more accurate results would be gained by using considerably smaller pixels. Use of already digitized Landsat MSS data, he notes, would have been preferable but could not be done given the economic constraints of this project. He concludes that logistic analysis is an ideal analytical tool for studies of this sort because it permits the researcher to incorporate variables of different levels (categorical and continuous) into the analysis. Finally, he points out that, although a nonprobabilistic archaeological sample of prior surveys was the basis of this experiment, a probabilistic sample would be more appropriate for future studies.

Another remote-sensing-aided predictive study in the western United States compared archaeological survey data from a 2.1 percent transect survey within the Naval Weapons Center at China Lake, California, with variables derived from resampled, 100 by 100 m pixel Landsat MSS data through a principal components clustering analysis of four-band data (Elston et al. 1983). The "major objective was to develop and characterize signatures for each transect irrespective of site content," and thus to arrive at an "independent typology of transects against which we can investigate the relationship between transect type and site occurrence" (Elston et al. 1983:63). The derived transect typology was displayed as a dendrogram, and the number of sites per surveyed transect were "superimposed on the distal nodes of the dendrogram" (Elston et al. 1983:64). The success of this projection was tested by arbitrarily selecting 45 more transects, classifying them according to their place on the dendrogram through additional Landsat-based cluster analysis, and surveying them to determine how faithfully the projection was borne out.

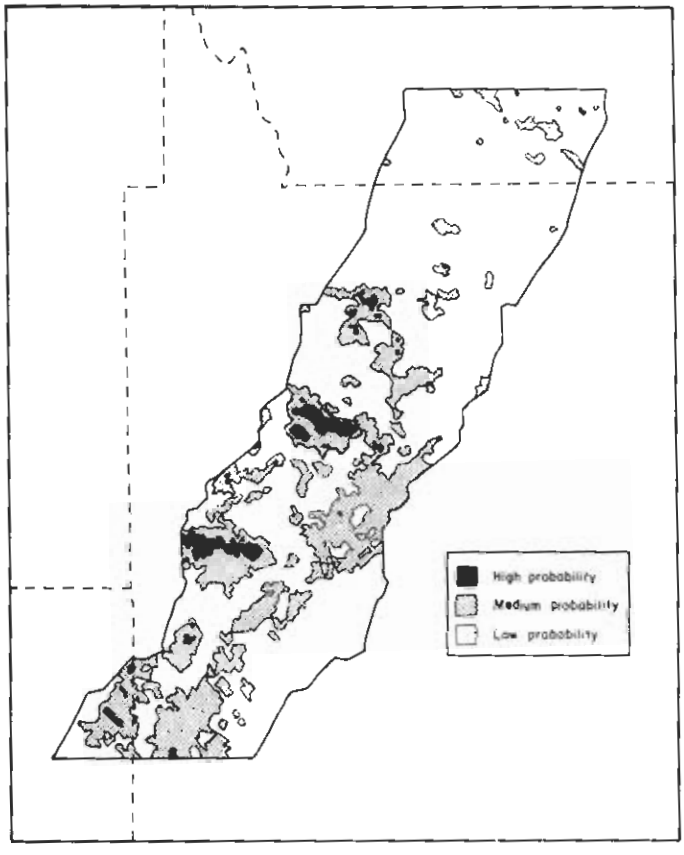


Figure 9.6. A site-occurrence probability map of a Colorado study area compiled for management purposes through digital analysis of Landsat MSS data (after Holmer 1982:Fig. 4.5.1). Darker shading indicates areas with high probability of site occurrence; lighter shading, areas with medium probability; no shading, areas with low probability.

Elston et al. found that their success rate for correctly characterizing the probability of site occurrence for transects was 86 percent. They suggest that the lower success rates of 58–70 percent achieved by Holmer (1982) and Baker and Sessions (1982) were a result of using a two-group—site-present vs site-absent—solution, when in reality not all "likely" areas would have been used in the past in sparsely occupied regions. Sites would also, according to Elston et al., be likely to occur or to not occur in more than one type of environmental setting. They characterize their approach as "natural" (Elston et al. 1983:66), not a "cookbook application of discriminant functions" like previous projects. The final results of their analyses were mapped in three transect classes: those with probabilities for site occurrence of less than 0.22, those with probabilities from 0.22 to 0.62, and those with probabilities from 0.62 to 0.67 (Figure 9.7), the last of which they term *site likely*. The narrow probability range represented by the group of site-likely transects is interesting and seems to indicate that there were a significant number of transects within this taxon.

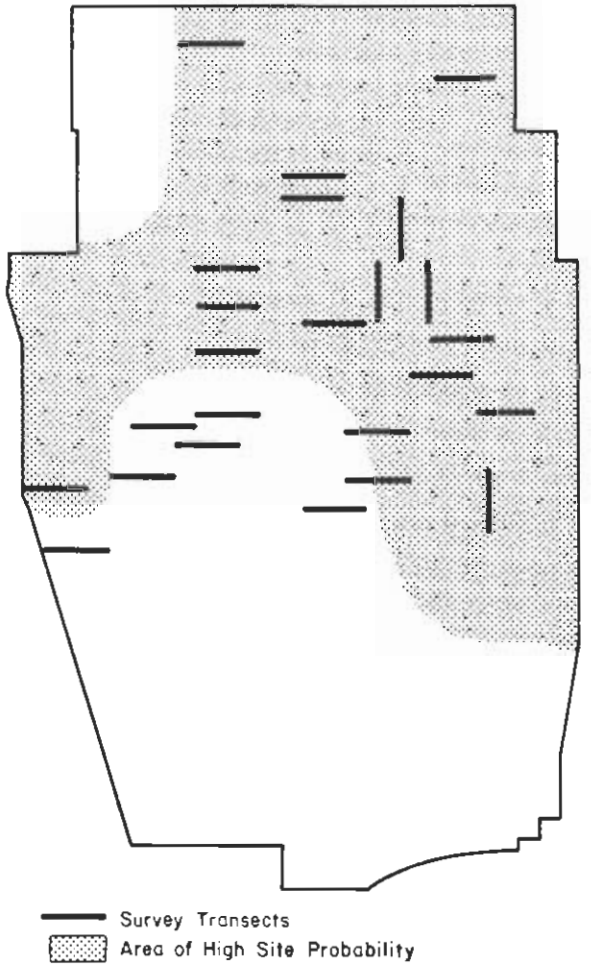


Figure 9.7. A site-occurrence probability map derived through digital analysis of Landsat MSS data at the Naval Weapons Center, China Lake, California (after Elston et al. 1981).

Another group of predictive archaeological experiments that made use of remote sensing for the measurement of environmental variables took place not in the arid West, but instead in heavily vegetated Delaware. Based on Wells's (1981) proposed model of the correlation between site locations and certain landform features (especially sand ridges), these predictive attempts have encompassed at least three separate archaeological studies (Wells et al. 1981; Custer et al. 1983; Custer et al. 1984). In each study, environmental variables included distance to water, geomorphological/landform setting, soil type, gradient, and convexity of the

landscape (the two latter variables presumably based on topographic and not Landsat data). These variables were measured within a 3500 m radius of each cell (Wells et al. 1981), and a training set of such measurements was used as input to a logistic regression comparing site occurrence with each variable. The study reported in Wells et al. (1981) resulted in the compilation of a general site occurrence probability map, which was then tested by further survey. This research indicated that the relative contribution of each variable to explaining site occurrence was as follows:

1. distance to minor stream, 50 percent;
2. distance to major stream, 42 percent;
3. distance to openland soils, 51 percent;
4. gradient, 51 percent;
5. convexity, 67 percent; and
6. distance to present marsh, 12 percent.

The low contribution of the last variable was explained by noting that most present marshes have been drained historically.

A second study (Custer et al. 1983) compared archaeological survey findings in the St. Johns and Murderkill drainages in Kent County, Delaware (Custer and Galasso 1983), on a period-by-period basis with Landsat-generated environmental variables, again using a logistic regression model and the same variables used in the previous study. Contour maps showing areas with less than 0.5, 0.5–0.75, and greater than 0.75 probabilities of containing sites were generated (Figure 9.8). During a second-stage test survey, 37 percent of the inspected areas that had been predicted to have probabilities in the 0.5–0.75 range contained sites, as did 49 percent of the surveyed areas with predicted probabilities of 0.75 or greater. It is not clear whether areas predicted to have less than a probability of 0.5 were tested.

Another, more comprehensive test of the Delaware models has only recently been reported. This study took place in New Castle and Kent counties as part of planning for a proposed highway corridor (Custer et al. 1984). Detailed *explanatory* site location models—i.e., theoretical formulations describing assumed past subsistence and mobility organization—were set forth for each temporal period prior to operationalization of the cultural and environmental variables. Environmental variables were then devised and measured using the University of Delaware's ERDAS 400 digital image analysis system. The authors then used their settlement pattern model to predict the distance to each of these landscape features from each site type. Using Wells's (1981) logistic regression method, Custer et al. produced contoured probability maps that again showed three probability zones of less than 0.5, 0.5–0.75, and greater than 0.75. These maps were compiled at 1:24,000 scale on 10 USGS topographic quads, and they are currently being used by the Delaware Department of Transportation as planning aids.

Although Landsat digital MSS data are the most likely source for the remote-sensing-aided classification and measurement of environmental variables, there

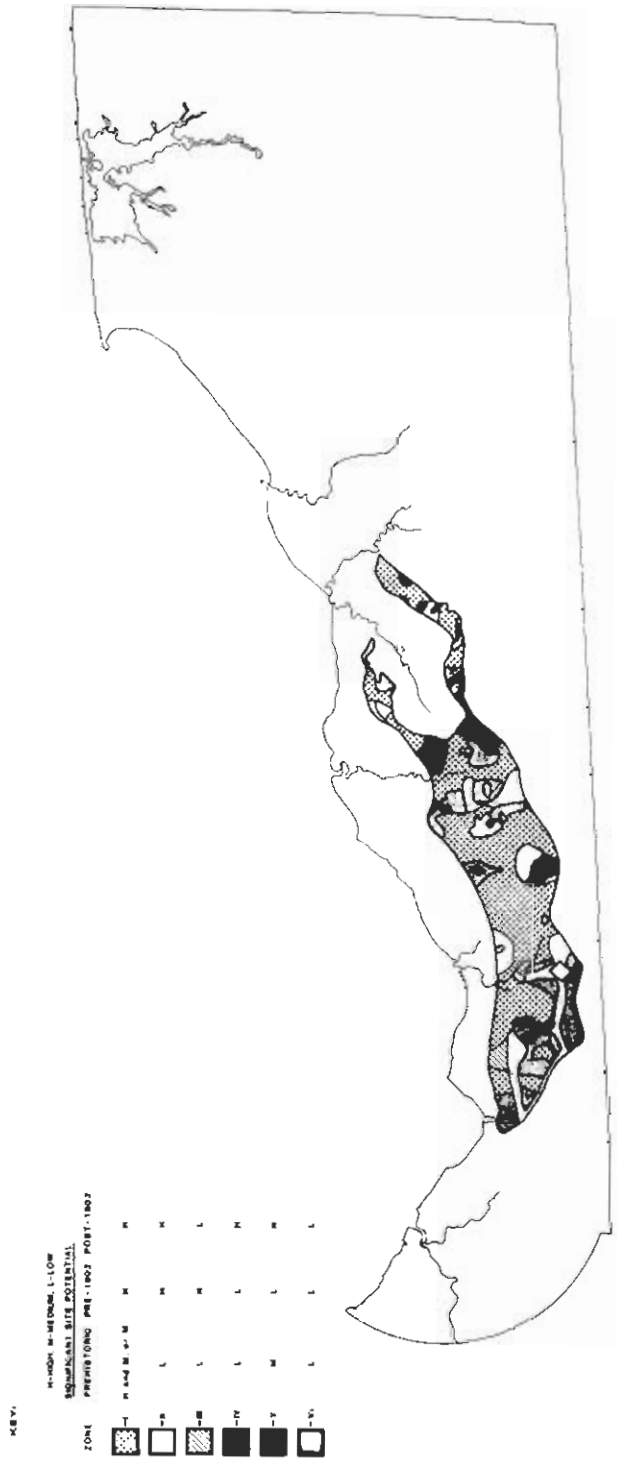


Figure 9.8. A probability map showing the potential for the occurrence of archaeological sites from six cultural/temporal periods in Delaware (after Custer et al. 1984:Fig. 26). Of all the archaeological "prediction" attempts known to the author, this study incorporates the most stringent modeling of cultural system dynamics and comes closest to actually predicting in an explanatory sense.

may be other choices in the near future. A recent experiment in assessing cultural resources in Bandelier National Monument in central New Mexico (Figure 9.9) made use of simulated SPOT digital data (Inglis et al. 1984). SPOT is a satellite that is soon to be launched by the Centre National d'Etudes Spatiales (CNES) in France; an airborne multispectral scanning device was flown by the CNES over selected targets in the United States so that scientists could experiment with the three-band SPOT data prior to launch. The Bandelier data were acquired on June 19, 1983, at a resolution of 20 m, more than twice the resolution of Landsat MSS data, and were analyzed using NASA ELAS software on a VAX 11/750 system at the University of New Mexico's Technology Applications Center.

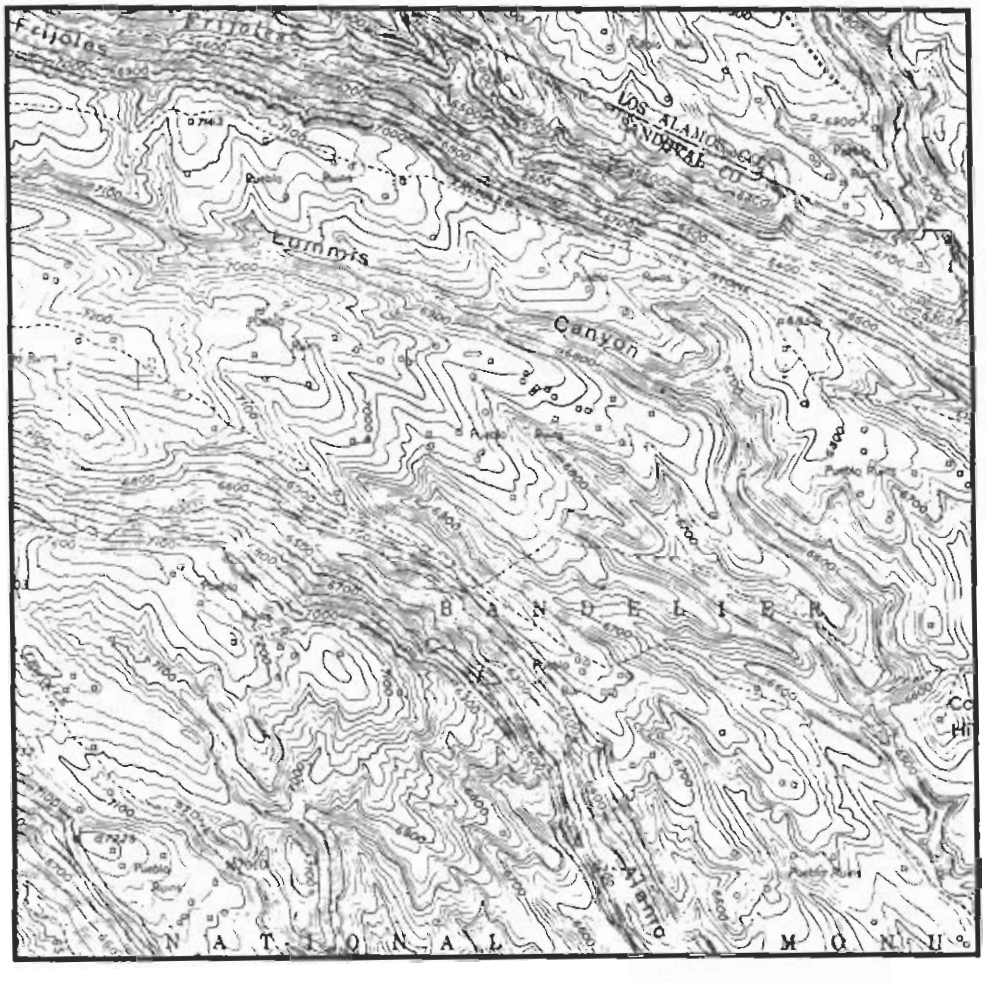


Figure 9.9. The location of an experiment in projecting archaeological site occurrence using simulated SPOT data in Bandelier National Monument, north-central New Mexico (after Inglis et al. 1984:Fig. 2; scale: 1:24,000, known site locations shown as open squares). Data from the French SPOT satellite will derive from a multispectral scanner with considerably higher resolution than that provided by Landsat. Although the data will be more expensive to acquire, they may be more cost-effective than Landsat data for cultural resource management.

The data were first geometrically corrected, then classified into 27 data classes by means of an unsupervised principal components analysis. The classification was output at a scale of 1:12,000 and superimposed on a map showing the locations of known archaeological sites in the study area. Two of the spectral classes, which together constituted 25 percent of the total study area, contained 45 percent of the known archaeological sites (Table 9.2); these classes do not correspond exactly with any previously mapped soil or vegetation classes in the monument, and the SPOT-generated projection is a more powerful indicator of archaeological site occurrence than any previous map. This suggests that SPOT or other high-resolution digital satellite data may have considerable potential for projecting the occurrence of archaeological materials in an empirical manner—or for predicting differential *visibility* of archaeological materials, as will be discussed at length below.

Remote-Sensing-Aided Archaeological Predictions: Some Comparisons and Comments

There are, of course, differences among the numerous studies in archaeological prediction summarized above. Some employed visual interpretation, while others were based upon computer analysis of digital data; some were approached as ways of designing samples using informed environmental stratification, while others were explicitly directed toward the "prediction" of archaeological site locations, occurrence vs nonoccurrence, or densities. The mathematical models used to compare dependent (cultural) variables with independent (environmental) variables vary as much in these remote-sensing-based approaches as they do in other types of "predictive modeling" that are not based on remote sensing.

These studies are basically the same in one sense, however. None of these attempts at prediction really constitutes prediction in the explanatory sense of the term advanced in Chapter 4. Each projects empirically from the known occurrence of archaeological sites to the probable occurrence of similar places in areas that have not yet been surveyed. This is a reflexive exercise, and it is somewhat unsatisfying in that there is no assurance that any such projection will be successful until it is tested, or that the next projection from the same data will be similarly successful or unsuccessful when tested. This is because, regardless of whether they incorporate a modern and useful technology like remote sensing, such nonexplanatory exercises do not focus on the systemic level of the explanation of past human organization. The next section of this chapter will set the stage for a discussion of some of the ways in which remote sensing might be used to produce more productive, explanatory models.

In concluding this section I would reiterate my caution that the use of Landsat and other remote sensor data should be carefully considered with regard to the limitations of this technology. Remote sensor data exist in the present and are no more "reflections of the past" than are contemporary archaeological data. Such landscape characteristics as the location of water and other factors change through

TABLE 9.2.

A 25 by 25 pixel (picture element) matrix of simulated SPOT data over part of the Bandelier National Monument study area. Known archaeological sites are shown as three-digit numbers in boldface. Known site locations were correlated with zones classified using cluster analysis of SPOT-simulated data (the single-digit numbers), and it was found that a majority of sites occurred in only a few zones.

	339	*338*	*339*	*340*	*341*	*342*	*343*	*344*	*345*	*346*	*347*	*348*	*349*	*350*	*351*	*352*	*353*	*354*	*355*	*356*	*357*	*358*	*359*	*360*	*361*
338	7	7	2	2	7	7	7	7	7	7	7	7	7	7	7	9	8	4	4	5	5	19	19	5	5
339	7	7	7	2	2	2	2	2	2	7	7	7	7	7	7	8	8	7	7	8	8	5	5	5	5
340	7	7	7	2	2	2	2	2	2	7	7	7	7	7	7	7	7	7	7	8	8	5	5	5	5
341	7	7	7	7	7	2	2	2	2	7	7	7	7	7	7	7	7	2	2	2	2	7	7	4	4
342	2	7	7	7	7	2	2	2	2	2	2	7	7	7	7	7	7	2	2	2	2	7	7	4	4
343	2	7	7	7	7	7	7	2	2	2	2	119	7	7	7	7	7	4	4	4	4	4	4	4	4
344	7	7	7	7	7	7	7	2	2	7	7	7	7	7	7	7	7	4	4	4	4	4	4	4	5
345	7	7	7	7	7	2	2	2	2	7	7	7	7	7	7	7	7	7	7	8	8	4	4	3	3
346	2	7	7	7	7	2	2	2	2	2	2	7	7	2	2	2	2	7	7	8	8	4	4	4	4
347	2	7	7	7	7	7	7	9	9	2	2	7	7	2	2	2	2	2	2	2	7	7	7	7	4
348	2	7	7	7	7	7	7	**101**	9	2	2	2	2	2	2	2	2	2	2	2	7	7	7	7	7
349	2	2	2	2	2	7	7	2	2	2	2	2	2	2	2	2	2	2	**112**	2	2	2	7	7	7
350	2	2	2	2	2	7	7	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	7	7	2
351	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	7	2	2	2	2	2	2	2
352	2	2	2	2	2	2	2	2	2	7	7	7	9	9	7	7	7	7	2	2	2	2	2	2	2
353	2	2	2	2	7	7	7	**114**	7	7	7	9	9	7	7	7	7	7	7	7	7	7	2	2	2
354	2	2	2	2	7	7	7	7	7	2	2	2	2	2	2	2	7	7	7	7	7	7	2	2	9
355	2	2	2	2	7	7	7	7	2	2	2	2	**111**	2	2	2	7	7	2	2	2	2	2	2	9
356	2	2	2	2	7	7	7	7	2	7	2	2	7	7	2	2	7	7	2	2	2	2	2	2	2
357	2	2	7	7	7	7	7	7	7	7	2	2	7	7	2	2	7	2	2	2	2	2	2	2	2
358	2	2	7	7	7	7	7	7	7	8	8	8	2	2	9	9	2	2	2	2	2	2	2	2	8
359	2	2	2	2	7	7	7	7	8	8	6	8	2	2	9	8	2	2	2	2	2	2	7	7	7
360	2	2	2	2	2	7	7	7	8	8	7	7	9	9	9	9	9	2	2	2	2	2	2	7	9
361	2	2	2	2	2	2	2	2	8	8	6	6	9	9	9	9	9	2	2	2	2	2	2	9	9
362	2	2	2	2	2	2	2	2	8	2	2	2	2	2	2	9	9	2	2	2	2	2	2	9	2

From Inglis et al. 1984:11

time, and there is no assurance that the reasons why activities were located where they were in the past will have any sort of transparent relationship to what we see on aerial photos or space images. In the final portions of this chapter some possibly more realistic ways in which contemporary remote sensing data and contemporary archaeological data can be brought to bear upon one another will be explored.

POTENTIAL APPLICATIONS OF REMOTE SENSING WITHIN THE EXPLANATORY FRAMEWORK OF ARCHAEOLOGICAL MODELING AND PREDICTION

So far in this chapter I have discussed remote sensing and what it is, what some of its general limitations in archaeology might be, and some of the ways in which archaeologists have applied remote sensing methods and data to experiments in predicting certain aspects of the archaeological tecord. Although remote sensing has been used in a number of different ways in these archaeological experiments, their general method is uniformly one of empirical, inductive "prediction" as diagrammed in Figure 4.1 of Chapter 4. This exercise generalizes from known distributions of archaeological sites or materials—known on the basis of prior surveys or the compilation of extant site forms—to a "prediction" of what additional sites or materials will be discovered in the future in areas not yet surveyed. This is accomplished through the tabulation or correlation of the differential occurrence of archaeological sites with respect to differential distributions of environmental charactetistics that are assumed to have been important to decisions about where sites would be placed. As discussed in Chapter 4, this is the method used in most of the "predictive modeling" efforts described in the archaeological literature or in management reports today; the limitations of and problems with this method are also explored at length in that chapter.

Chapter 4 also describes another way of thinking about modeling and prediction—as integral aspects of the process of archaeological explanation. Referring again to Figure 4.1, the interpretations we make concerning the archaeological record (that is, the meaning we assign to the remains that we encounter) are separated from the actual physical nature of the archaeological record by many levels of phenomena.

It is the physical archaeological record and its distribution that managers are interested in, for this is what they must manage. Meaning is given to the archaeological record, however, only through explanation, and meaning is essential to predictive or projective statements about the physical archaeological record for two ptimary reasons. The first is that in order to predict the locations of archaeological sites successfully we must know not only what "noncultural variables" they are correlated with, but also *why*. The answers to the questions "why?" must be posed in terms of systemic human organization, because systemic organization is the way that people differentially locate themselves and their activities on a landscape. If we

do not understand the *systemic mechanisms* of site placement, then there is no assurance that any prediction can be extended from a known area to unknown areas, even if these areas are immediately adjacent to one another. Such mechanisms can be known only through explanatory modeling, and through use of predictions to test these models.

In Chapter 4 Kohler and I have suggested that the best independent variables for site location models are *ecosystemic* variables. Before models that incorporate ecosystemic variables can be formulated, however, the things that intervene in the real world between the physical archaeological record and past human organization must be addressed or "filtered out." These are the factors listed in the "processes" column of Figure 4.1: discard behavior, depositional and postdepositional processes, and the methods that archaeologists use to discover, measure, and analyze the portions of the archaeological record that we find.

The other reason that the meaning given to the archaeological record—explanation—is all-important in managing this record is easier to state but ultimately more difficult to define. As the volume editors have pointed out in Chapter 1, the legal and, I like to think, moral reasons for even worrying about managing cultural resources are based on the significance of those resources in terms of research potential. Cultural resources are important because, by using them, archaeologists may be able to say something worthwhile about the operation and organization of human systems and their components, past *and* present. The management of *significant* archaeological resources has been mandated, and significance is based on meaning given to cultural resources through the explanatory framework of archaeological science.

I can suggest two ways in which remote sensing has the potential for moving archaeological prediction away from simple empirical generalization and toward more explanatory goals. It should be understood that remote sensing, while it can play a part in this reorientation of "predictive modeling," is not the solution in itself. The real solution lies in the ability of archaeologists to change the ways in which they think about doing archaeology—particularly we must discard the idea that archaeological explanation is or can ever be *easy*. Remote sensing can only play a part in shifting archaeological thinking, but this part may be indispensable because of the unique and inclusive sorts of data that remote sensing can provide. Remote sensing can provide two specific and immediate classes of data that archaeologists need: data pertinent to depositional and postdepositional processes, and data through which ecosystemic, rather than simply environmental, variables might be measured. A few experiments in measuring and using such data are reported below, along with suggestions concerning possible future directions.

Another area in which remote sensing can aid in the investigation and explanation of the organization of past human systems is through applications to ethnography and ethnoarchaeology. As emphasized previously in this chapter, remote sensor data are contemporary, and as such they might best be applied to understanding the relationships between ongoing hunter-gatherer and primitive agricultural systems. These relationships are one of the most exciting data sources for

archaeologists attempting to understand the operation of past systems because they have the potential for suggesting some of the alternative ways that people adapt to differing conditions through time and across space. Some of the ways that such information can be brought to bear on the question of predictive modeling are suggested by Ebert and Lyons (1983), Kruckman (1972), Parrington (1983), and Vogt (1974).

Remote Sensing and the Measurement of Depositional and Postdepositional Processes

The materials that people use leave the cultural context and enter the archaeological context when they are discarded; at some point after they are dropped on or intentionally buried under the surface of the earth, they come under the influence of depositional processes and are incorporated in sediments and soils. Deposition most often occurs in the context of aggradational processes that bury cultural materials, although there are situations in which cultural materials remain on the surface of the ground. Some depositional processes are cultural, consisting of burial by human activity, but these are less common than natural depositional events.

Materials buried in a definable layer or "level" are often assumed to be the results of a single occupational episode (Conkey 1980), but this is not necessarily always the case. The nature of the deposited archaeological record is controlled by the periodicity of occupation or use of a place and the relationship between this periodicity and the periodicity of depositional processes acting on cultural materials. Artifacts that are dropped only sporadically might be covered by sediments left by depositional processes that occur more often than episodes of dropping, while artifacts that are lost or abandoned relatively continuously will often be subjected to depositional processes only after several episodes of site occupation have taken place. In the latter case, the apparent "levels" will be the result of more than one episode of site use. For instance, if a site is occupied or is the locus of activity several times between successive rainy seasons, more than one episode of activity may be represented in each depositional level. This poses problems for the archaeologist who is attempting to sort out the results of periodic human behavior in that "demonstrably associated things may never have occurred together as an organized body of material during any given occupation" (Binford 1982:17–18).

Once cultural materials are deposited and become part of the archaeological record, they are acted upon by another set of processes that can be thought of as postdepositional. Most processes that disturb or act upon the surface or subsurface of the earth also affect archaeological deposits. Such biological processes as faunalturbation and floralturbation (Wood and Johnson 1978) modify deposited materials, as do a host of other mechanical and chemical events. Foley (1981) presents a taxonomy of natural processes responsible for the burial, movement, destruction, and modification of archaeological deposits (reproduced here as Figure 9.10). Dis-

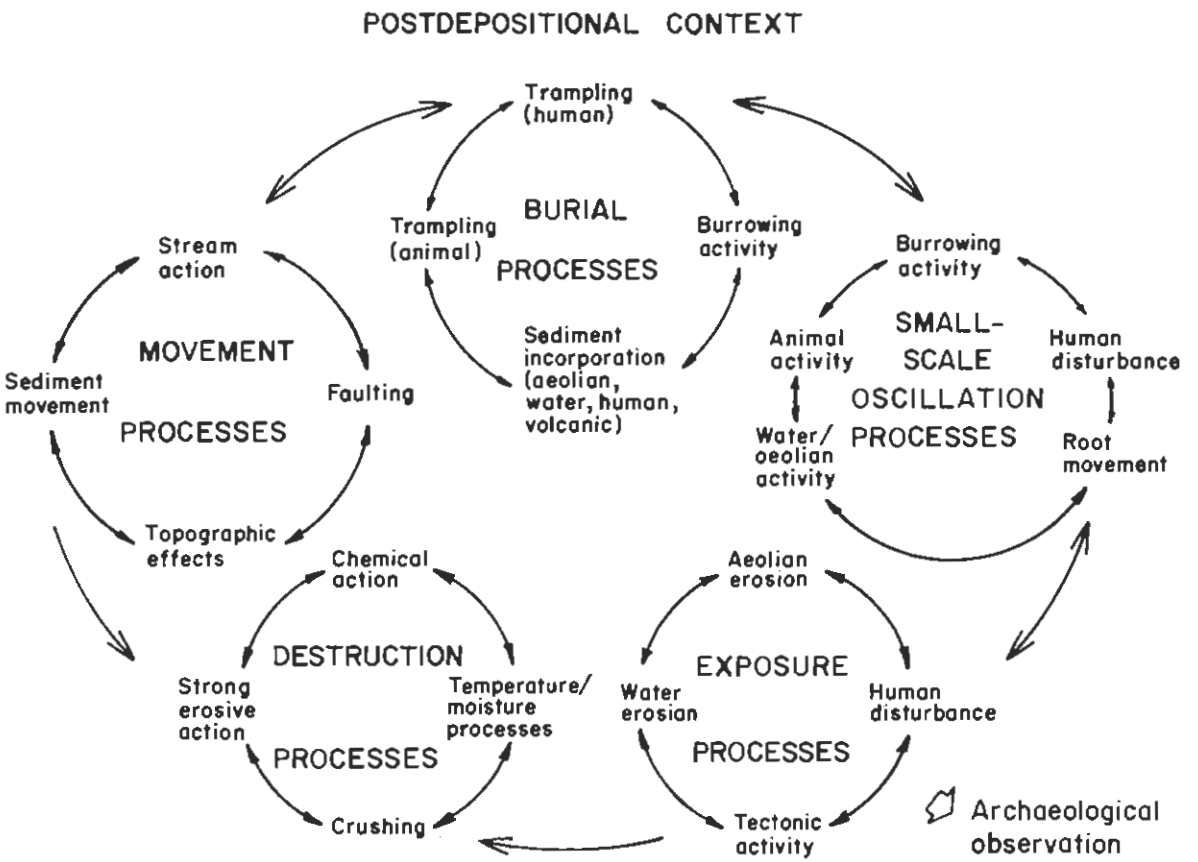


Figure 9.10. Cultural and noncultural postdepositional processes are probably responsible for much of what we, as archaeologists and managers, find during archaeological survey. This taxonomy of such processes (after Foley 1981) illustrates one way to think about the complexities of formation processes of the archaeological record. All such processes must be taken into account before we can arrive at true explanatory predictions using remote sensing or other data collection methods.

carded materials enter the archaeological record through burial by cultural or natural agencies; once assemblages are buried they may remain in place, or they may be moved through stream action, sediment movement, faulting, or mass wasting. At the same time, certain materials may or may nor be altered by physical and chemical agencies while in or on the ground. Foley (1981) also identifies what he calls small-scale oscillation processes that act on the discarded archaeological record, including water or wind action, animal burrowing, root action, and human disturbances.

Natural postdepositional processes can alter or destroy archaeological materials, but they also play a role that is vitally important to the archaeologist: they expose these materials, making them visible and thus available for study. Most archaeology carried out in the United States today is undertaken in the context of cultural resource management assessments, which entail systematic survey of the surface of the earth in areas that are to be disturbed by reservoir construction, strip mining, or other engineering and resource-extraction activities. Buried archaeological materials are not found during such surveys; only those cultural materials that are exposed but not totally destroyed are found and serve as the basis of archaeological study and interpretation. When subsurface testing is incorporated into surveys, it can expose but a tiny part of buried remains. It is only during the short and relatively uncommon period between the exposure of deposited materials and their dispersal or destruction that these materials are available to archaeologists for study. For this reason, it is critical that archaeologists carefully consider the nature and actions of the processes that make their basic data available to them.

There is no easy way for the archaeologist to observe, characterize, measure, and predict depositional and postdepositional processes. Both deposition and most postdepositional alteration took place in the past, so these processes cannot be observed directly. In addition, the distribution of these processes probably varies across the landscape. Analogs might be found in contemporary surface processes, however, which means that the forces that have acted on archaeological materials (and possibly also their rates or the magnitude of their effects on the archaeological record) are potentially predictable. If such processes can be predicted, then at least some aspects of the depositional and postdepositional "formation processes" (Schiffer 1983:675) intervening between the materials discarded by past peoples and the archaeological record that we actually see today can be taken into account. And such factors must be accounted for before we can attempt to predict the locations in which archaeological materials can be expected.

To most archaeologists it seems reasonable to turn to geologists and geomorphologists for the details of such natural processes and of their differential occurrence and rates, but usually these disciplines cannot provide the necessary level of detail. In fact, when an archaeologist and a geomorphologist are introduced, the latter will almost always initiate probing questions about whether archaeology can supply concrete dates for recent natural surface events. This interest on the part of geomorphologists has probably been the major impetus behind the development of the subfield of geoarchaeology (Butzer 1977; Gladfelter 1981), but it is just the

reverse of what we want to hear. Most geomorphological studies are conducted in circumscribed places under specific conditions and are even more inductively based than archaeology. Archaeologists need to be able to arrive at generalizations about the places in which different surface processes act to deposit and disarrange or preserve archaeological materials across relatively large study areas. Fortunately, remote sensor data, with their wide areal coverage, may help to supply this information.

One such remote sensing study was undertaken in an attempt to define the extent of different surface deposits and their archaeological correlates in Chaco Canyon in northwestern New Mexico (Ebert and Gutierrez 1981). Chaco Culture National Historical Park has been extensively surveyed for at least 50 years owing to the spectacular and concentrated nature of its archaeology, and a data base of more than 1200 archaeological sites was available at the National Park Service's Division of Cultural Research for comparison with remote-sensing-aided mapping of surface deposits there. Previous geological and geomorphological studies had examined alluvial deposits and hillslope processes and their rates, and these data provided a basis for photointerpretation and mapping of geomorphic surface units.

Geomorphic units were interpreted by Ebert and Gutierrez (1981) using 1:6000 scale aerial color transparency photos viewed with a Bausch and Lomb variable-power stereoscope; these units were transferred to 1:12,000 black-and-white orthophotoquads and from those to a 1:12,000 scale base map, which also bore the locations of archaeological sites in the data base. Two descriptions—landform and photointerpretive—were generated for each geomorphic unit defined, based on tone, color, texture, vegetation associations, and landform associations (Figure 9.11 and Table 9.3).

Correlations between site locations and geomorphic surface units (summarized in Table 9.4) were of interest relative to interpretations of the differences between locations where different types of sites were found by survey archaeologists. Archaic sites, usually consisting of small scatters of stone flakes, were found on the oldest visible surfaces in Chaco Canyon. Similarly, Basketmaker sites were found primarily on stable and inactive surfaces, as were the Pueblo I, II, and III sites. Later Pueblo sites were found relatively more often on less stable surfaces, and the even more recent Navajo sites occur in high proportions on very active surfaces where older materials would either be obscured or destroyed. The smallest sites (as recorded in the NPS data base) are found in units with little or no alluvial or aeolian surface veneer, while larger sites predominate in fine-grained, inactive Quaternary units where sheetwash, uniform sedimentation, and relatively even aeolian deposition would cover smaller occurrences but allow larger materials (masonry walls, for instance) to project above the surface.

Another remote-sensing-based study, which built upon the Ebert and Gutierrez (1981) Chaco Canyon experiment, was carried out in the Green River Basin of southwestern Wyoming (Wandsnider and Ebert 1983). Fluvial, aeolian, and gravitational processes have altered the landscape there in post-Pleistocene times, giving rise to what appears to be a varied and diverse region when it is considered on a

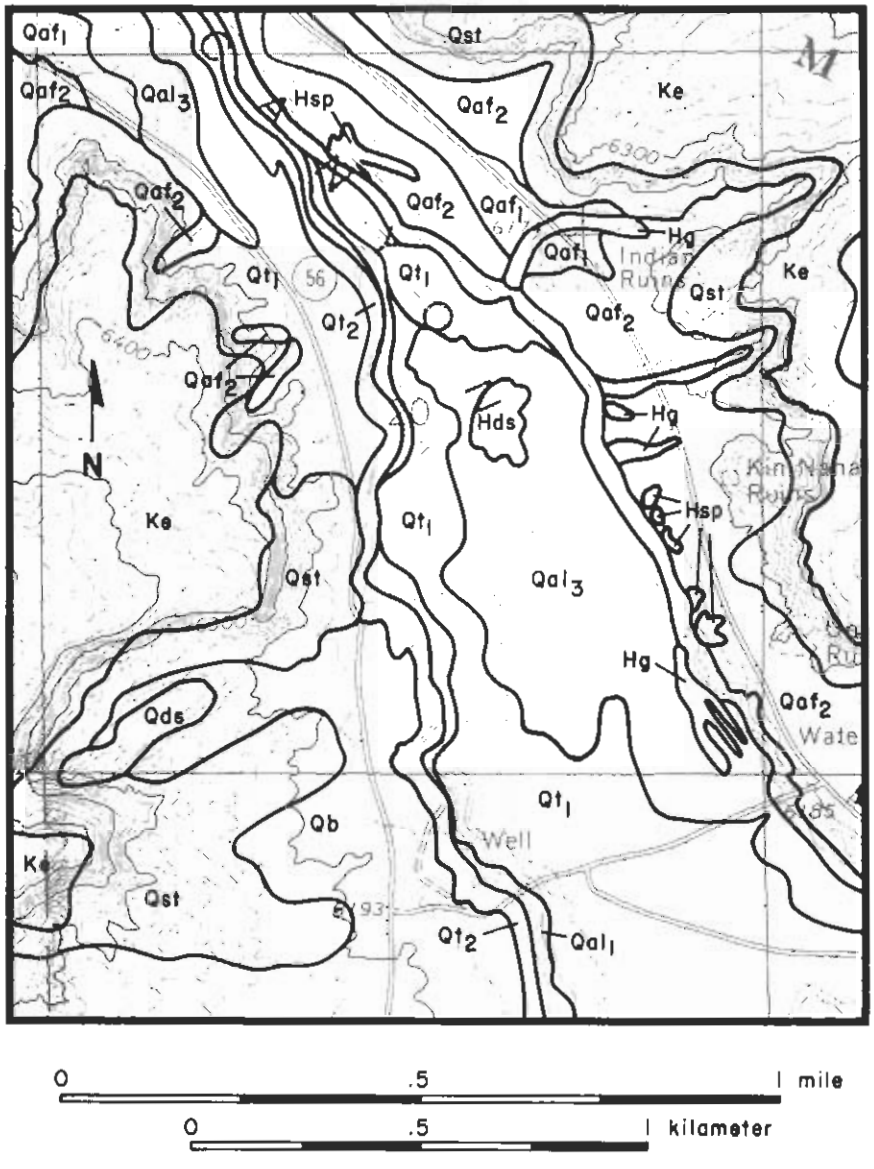


Figure 9.11. An experiment carried out by the National Park Service's Branch of Remote Sensing involved the mapping of geomorphological surface units at Chaco Culture National Historical Park to explore the role of postdepositional processes affecting the visibility, integrity, and discovery of the archaeological record. This is a portion of the map of surface units that were derived using photointerpretation of 1:12,000 scale color aerial photographs at Chaco Canyon in northwestern New Mexico (after Ebert and Gutierrez 1981:Fig. 1). Descriptions of the various units appear in Table 9.3.

TABLE 9.3.

Surface geomorphological units, their designations, their photointerpretive recognition patterns, and their descriptions and summary surface dynamics as mapped using photointerpretive techniques in the study of postdepositional processes on the archaeological record at Chaco Culture National Historical Park in northwestern New Mexico

Map Unit	*Designation*	*Landform/Photo Description*	*Stability/Dominant Processes*
$QT_{p(s)}$	Quaternary Tertiary pediment deposits (stable)	Undulating surface with poorly integrated surface drainages, dark in tone. Vegetation sage and grass.	Stable alluvial, colluvial, and aeolian deposits resting unconformably on eroded Tertiary and Cretaceous deposits. Little runoff or sediment produced on these highly permeable deposits.
$QT_{p(t)}$	Quaternary Tertiary pediment deposits (transitional)	Tan sand texture, scattered vegetation and integrated surface drainages. High drainage density, parallel to dendritic drainage pattern.	Produces significant runoff and high sediment yields; occupies zones between $QT_{p(s)}$ mesas and unvegetated badlands.
Kb	Cretaceous sandy bedrock (buried)	Light tan-white bedrock with discontinuous veneer of aeolian sand, local sheetwash alluvium.	Intermittent aggradation/erosion by sheetwash, aeolian processes.
Ke	Cretaceons sandy bedrock (exposed)	Light tan-white exposed bedrock with very sparse vegetation; fine textured with joint patterns clearly visible. Cliff House and Picture Rocks formations.	Flat surface with little or no cover; sheetwash and aeolian erosion.
Hg	Holocene gullies	Localized discontinuous drainage, 1-3 m deep, up to 300 m long.	Unstable, rapidly eroding.
Hds	Holocene dam sedimentation	Well-vegetated, fan-shaped deposits behind dams or diversions.	Rapidly aggrading, anastomosing channels; date from 1930s.
Hsp	Holocene soil pipes	Arcuate depressions or collapsed soil pipes on terrace edges.	Soil piping, mass movement, highly unstable, eroding rapidly.
Qaf_1	Quaternary alluvial fan 1	Topographically raised, irregular-shaped deposits; vegetation density slightly higher than Qaf_2 or Qt_1.	May or may not contain active, incised channels.
Qaf_2	Quaternary alluvial fan 2	Conical fan-shaped fill associated with major side canyons; light-medium tone.	Relatively stable surface, some graded to QT_1 surface. May contain buried soils of humus-rich layers.

Qal_1	Quaternary alluvium 1	White, incised, meandering/straight channel in central valley with vertical walls.	Active alluvium, thickness 0–1.5 m. Erosion/aggradation dominant.
Qal_2	Quaternary alluvium 2	Tan–light gray alluvium in incised channels with steep banks, cutting through alluvial fill in side.	Active alluvium in major tributaries of Chaco Canyon, many individual channels and cut/fill sequences.
Qal_3	Quaternary alluvium 3	Light-toned alluvium associated with surface traces of fossil channels, high vegetation.	Inactive alluvium, mostly reworked Qal_2 material. Thickness 2–4 m.
Qb	Quaternary badlands	Banded gray to dark brown beds following topography, high drainage density. Little vegetation.	Relatively impermeable shales with interbedded sandstones; covered by 0–0.7 m of weathered mantle. Easily eroded, active surface.
Qc	Quaternary fine-grained colluvium	Light brown, fine textured, irregular shaped deposits near extensive shale and sandstone outcrops.	Sheetwash material derived from valley sidewall sandstones and gentler shale slopes at their base.
Qds	Quaternary dune sand	Light brown, linear, topographically high deposits associated with Ke bedrock mesas and buttes. Bushes; no grasses; no established drainages.	No integrated drainage development, little erosion. Dunes aligned N 60–70° E where linear. Thickness 0–2 m.
Qst	Quaternary talus	Medium-tone bands along base of sandy bedrock cliffs. Large angular blocks of sandstone talus on shale slope.	Larger talus blocks stable, localized creep, sheetwash, and debris flow deposits.
Qt_1	Quaternary terrace 1	Highest terrace incised by current Chaco Arroyo (Qal_1). Large areas of low relief within main canyon. Vegetation sparse.	Oldest inactive terrace; interbedded with alluvial fan, sheetwash, and colluvium from side canyons.
Qt_2	Quaternary terrace 2	Discontinuous, light brown, fine textured areas between Qt_1 scarp and active arroyo (Qal_1).	Youngest terrace or floodplain of present arroyo (Qal_1) in some areas. Stability varies, 1–3 m above channel.

From Ebert and Gutierrez 1981

TABLE 9.4.

Occurrence of known archaeological sites and materials at Chaco Culture National Historical Park (grouped by cultural affiliation and site size) within geomorphic surface units mapped with photointerpretive techniques

	Geomorphic Unit																
	Hg	*Hsp*	*Kb*	*Ke*	Qaf_1	Qaf_2	Qal_1	Qal_2	Qal_3	*Qb*	*Qc*	*Qds*	*Qst*	Qt_1	Qt_2	$QT_{p(s)}$	$QT_{p(t)}$
								CULTURAL AFFILIATION									
Archaic	0	0	.03	.05	0	0	0	0	0	0	0	0	0	0	0	0	0
Basketmaker II	0	0	.03	.03	0	0	0	0	0	0	0	.05	0	0	0	.06	0
Basketmaker III	0	0	.08	.10	.06	.11	.10	.07	0	.10	.09	.15	.05	0	.19	.06	0
Pueblo I	.20	0	.06	.14	.13	.22	.20	.31	.11	.27	.32	.30	.20	.17	.29	.19	.06
Pueblo II	.60	.50	.12	.27	0	.22	.30	.36	.56	.28	.29	.35	.28	.24	.24	.31	.37
Pueblo III	.20	0	.11	.22	.25	.22	.30	.13	.22	.17	.12	.15	.22	.25	.14	.31	.32
Pueblo IV	0	0	0	.01	0	0	0	0	0	0	0	0	0	0	0	0	0
Navajo	0	0	.33	.18	.44	.20	0	.07	0	.15	.15	0	.19	.18	.14	.06	.14
Unknown	0	.50	.25	.01	.13	.03	0	.07	.11	.03	.03	0	.06	.16	0	0	.11
								SITE SIZE									
Very small	.25	1.00	.47	.40	.36	.33	.25	.43	.08	.26	.15	.19	.25	—	.35	.33	.40
Small	0	0	.21	.26	.43	.24	.50	.18	.23	.34	.40	.38	.26	—	.41	.67	.21
Medium	.50	0	.25	.22	.31	.35	0	.32	.54	.28	.30	.31	.39	—	.12	0	.29
Large	0	0	.04	.09	0	.04	0	.07	.08	.07	.15	.06	.05	—	0	0	.05
Very large	.25	0	.02	.03	0	.04	.25	0	.08	.05	0	.06	.05	—	0	0	.05
Sample Size	5	2	156	325	16	88	10	45	9	359	34	20	389	124	21	16	79

From Ebert and Guttierez 1981

small scale. The Green River Basin is quite arid today and probably has been for some time; in most places rainfall is less than 400–500 mm annually. Even on high plateaus and slopes, vegetation is sparse, usually covering not more than 20 percent of the ground surface, which makes a remote sensing approach to surface depositional, erosional, and aggradational processes fairly straightforward.

The Green River Basin seems to have been inhabited at relatively low population levels since the beginnings of North American settlement. Paleoindian, Archaic, Fremont, and Plains Indian groups have left their remains there for at least 10,000 years: it appears that there may actually have been little difference in the lifeways of these people over a long time span, although the Fremont were at least partially agricultural while the others followed a hunting and gathering way of life. The majority of archaeological sites found in the Green River Basin are Archaic, a broad typological category encompassing virtually all materials dating from about 9000 BC to historical times, with assemblages consisting of stone tools and debris and containing little or no pottery. Many "sites" found in the Green River Basin are hundreds of meters long and wide, contain tens to hundreds of hearths, and have relatively sparse but even distributions of lithic artifacts. These assemblages and features are very likely the result of the reoccupation of these places over many thousands of years, coupled with depositional and erosional processes encouraging the formation of superimposed assemblages or palimpsests.

The Green River Basin experiment coupled the mapping of natural surface processes with an on-the-ground archaeological survey carried out by the National Park Service Branch of Remote Sensing in 1983–1984. This experiment was directed toward evaluation of the cultural resources on lands surrounding the Seedskadee National Wildlife Refuge along the Green River that are under the jurisdiction of the Bureau of Reclamation. The explicit goal was to incorporate remote sensor data into a predictive model of archaeological site locations and their characteristics.

Before the zones of differential geomorphic surface processes affecting the archaeological record in a 559,000 ha (1,380,700 acre) study area could be mapped, a data source was needed that would provide a regional perspective while permitting discrimination of different sorts of areas with resolution at culturally and archaeologically relevant scales. Remote sensor data, particularly those derived from satellite-borne sensors, are ideal for this application, particularly where little on-the-ground geomorphological mapping has taken place. The basic data source used in geomorphological mapping of the project area was a 1:100,000 scale Landsat 3 color composite visual product. Composed of an overlay of bands 4, 5 and 7 data from the Landsat multispectral scanner, this image has a ground resolution of about 80 by 80 m and approximates a color infrared view of the imaged scene. Color infrared accentuates vigorous vegetation, permitting discrimination between areas of growing plant cover and bare earth; this capability is particularly useful in defining differential surface processes.

Mapping was initiated by overlaying a sheet of frosted mylar on the 1:100,000 scale Landsat scene of the study area and placing these two registered sheets on a light table. Black-and-white photo prints at a scale of 1:80,000 and arranged in a

mosaic fashion were checked against the Landsat image to define the boundaries of zones of different geomorphic processes on the mylat overlay. Geology maps prepared by the State of Wyoming and Soil Conservation Service provisional county soils maps for Sweetwater, Lincoln, and Uinta counties were also consulted during the interpretation process.

Although the information available from the Landsat image, the aerial photographs, and the maps was different, the three sources were found to be complementary. The resolution of aerial photographs is many times greater than Landsat resolution, of course, and permits identification of small-scale topographic patterning. For instance, individual sand dunes and interdunal flats could be easily distinguished on the aerial photographs. Once areas characterized by dunes were located on the aerial photographs, the same areas were checked on the Landsat image, and the tonal and textural qualities of those areas were noted. By using the patterns identified in this way, we were able to detect additional dune areas directly from the Landsat image, subject to verification using the aerial photographs after such an interpretation was made. In some cases the geological and soils maps were useful in checking and placement of boundaries, although these maps were far more generalized than the geomorphological mapping done from the Landsat data. Photointerpretation could have been performed using only the aerial photographs, but this would have required the construction of a control network (see Ebert 1984) for about 100 prints, a very difficult task. Landsat data are geometrically corrected; thus, these data are ideal for environmental mapping such as that undertaken in the Seedskadee project area.

Fifteen of the larger geomorphological zones (Figure 9.12) were grouped, for purposes of discussion, under six general headings with assumed depositional and postdepositional significance:

1. **Terraces** formed largely by fluvial processes. This class includes both presently active terraces and those formed in the more or less recent past—possibly as early as the Pleistocene. In the most recently active of these areas, channel and overbank deposition dominate the depositional processes, while on earlier terrace surfaces slopewash, sheetwash, and aeolian processes are common.

2. **Playas and Flats** consisting of relatively flat areas experiencing slow deposition of fine-grained sediments. Deposition in these areas is facilitated by either internal or external drainages. When dry, these areas are subject to aeolian deflation.

3. **Dunes**, which in the study area occur not in extensive fields but rather interspersed throughout badlands, flats, or along the edges of intermittent watercourses where sand is plentiful. Some dunes also occur where mesatop scarps cause the wind to drop its sediment load. Presently active dune areas, which form the majority of the areas included in this category, are characterized by connected crescentic or barchan dunes; at least two areas of earlier, relatively well stabilized parabolic dunes are also found in the study area.

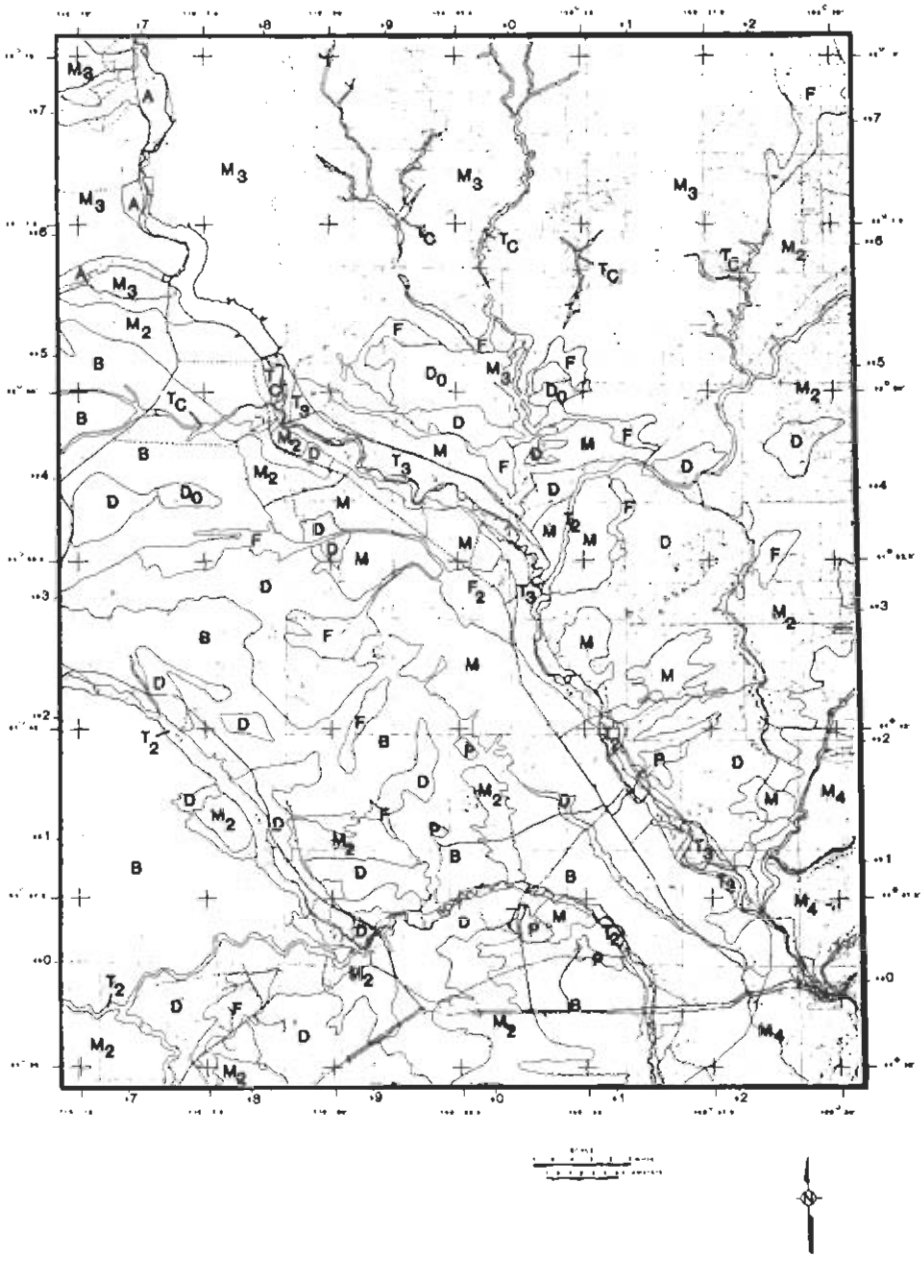


Figure 9.12. Geomorphological surface units interpreted visually from a Landsat color composite print and 1:80,000 scale aerial photographs in and around the Seedskadee National Wildlife Refuge on the Green River, southwestern Wyoming. This map was compiled as part of a distributional archaeological survey of the area (Wandsnider and Ebert 1983). Results indicate that much of what we see and bound as "sites" may be the result of relatively local differences in postdepositional processes and also is probably heavily influenced by the survey and recording methods employed. The "site" is not a concretely defined entity and may be an inappropriate unit of recording and analysis for purposes of explanatory archaeology.

4. **Badlands** consisting of highly eroded shales with a dense and reticulate drainage. In most cases this geomorphological class is interspersed with flats, dunes, and small remnants of earlier surface or "mesa" areas.

5. **Mesatop** areas, which are the more or less dissected remnants of earlier Tertiary gravel/sand bedrock surfaces. Four mesatop areas were distinguished on the basis of their landform and by the fact that at least some vegetation cover (dominated by Basin big sagebrush [*Artemisia tridentata tridentata*] and grasses) was distinguishable on the Landsat image of these areas.

6. **Agricultural Areas** irrigated with water from the Fontenelle Reservoir or the Green and Black's Ford rivers, which are extensively modified and probably need not be further considered by archaeologists, at least by archaeologists searching for surface remains.

Archaeological data were collected in this study area through a nonsite or distributional archaeological survey strategy (described at length in Chapter 4) to test these formulations and are still being analyzed. One pertinent observation made during the collection of the archaeological data was that the scale of surface processes with apparent relevance to artifact distributions may be far smaller than the scale of surface processes that can be discerned on Landsat MSS or small-scale aerial photographs. More recently, surface geomorphological processes have been reinterpreted using stereoscopic photointerpretation of 1:12,000 black-and-white aerial photographs of the 500 by 500 m sample units within which field archaeological recording took place. While the initial, small-scale photointerpretation was directed toward understanding *general* postdepositional characteristics across the study area, this second analysis will be applied directly to the task of filtering out postdepositional processes affecting specific archaeological materials found during survey. In order for this to be accomplished, it is clear that artifacts rather than sites must be the unit of discovery and recording. See Chapter 4 for a discussion of the advantages (and, I would suggest, the necessity) of a distributional archaeological approach.

Remote Sensing and the Measurement and Meaning of Ecosystemic Variables for Archaeological Modeling and Prediction

In Chapter 4 of this volume, Kohler and I have suggested that one avenue by which archaeologists might move beyond the empirical, inductive generalizations that we currently refer to as "predictive modeling" is by attempting to use ecosystemic rather than simply environmental or landscape characteristics as independent variables. It is the organization of human *systems* that we must understand if we are to explain the mechanisms behind mobility, the placement of activities in space, and the locations of discarded archaeological evidence. It was pointed out that at the systems level human organization responds not to the unique placement of specific resources at a single time and place, but rather to the regional spatial and

temporal patterning of all resources—that is, to the organization of the ecosystem as a whole.

There are abundant means for measuring simple environmental variables—slope angle and aspect, distance to water sources, elevation, and the like—and this is probably the major reason why these quantities are used as variables in most contemporary predictive models. Measuring or even identifying ecosystem variables is more difficult, and the first step in using such variables in modeling, prediction, and explanation will consist of research into new measurement techniques. Remote sensing is one source of such techniques that is increasingly available to the archaeologist. An example of a remote-sensing-based approach to the measurement of one possible ecosystem variable—environmental diversity—will serve as an illustration of possible research directions.

Environmental diversity, as the term is used here, is a measure of spatial heterogeneity in resources; even in a very general sense it is obvious that this variable should have consequences for the organization of human subsistence behavior. In an environment where many different resource species are distributed evenly, a human group dependent on these resources should minimize energy expenditure by being sedentary and territorial; if resources are clumped rather than evenly distributed, then high mobility will be necessary in order to exploit the full range of resources.

In order to examine the potential of this variable for explaining differences in human mobility and resource procurement, Harpending and Davis (1977:276) have suggested a "model" consisting of a one-dimensional environment along which the occurrence of a variety of natural resources is measured and for which the abundance of each resource is graphed as a continuous function. The complex continuous function represented by each resource can be viewed as the sum of Fourier components—a series of sine waves of different frequencies added together—and the resulting power spectrum can be analyzed.

Harpending and Davis initiate their model from the stance that hunter-gatherer groups seek or desire maximum variety in their diet, an assumption that is far from proven but one that is common in the Bushman literature and in fact in most literature dealing with generalist hunter-gatherers. If this assumption is correct, however, it is clear that people pursuing such an adaptation would seek areas in which to live and gather foods that had the maximum possible variety of food.

Harpending and Davis also hypothesize that the benefit that hunter-gatherers derive from increasing the size of their range is greatest when resources are out of phase—that is, they do not co-occur perfectly—with a cycle of redundancy of 1 km to 100 km. When all resources occur together at discrete locations, the benefit from increasing range size should be less. Maximum range size would be expected where there are few resources and where those resources are maximally out of phase with one another over distances of 1–100 km; minimum range size should occur where resources show little spatial variation or where many resources co-occur. Harpend-

ing and Davis also suggest some implications for group sizes: groups with maximum range sizes and extremely high mobility in low-abundance, out-of-phase resource environments should be relatively small with poorly defined local boundaries (for instance, in the Kalahari Desert). In the minimum range-size category, small groups would be expected with little spatial resource variation (e.g., in tropical rainforests), while larger groups would occur when resource variances are in phase (for example, on the northwest coast of North America).

A test of anthropological and archaeological implications of such expectations would depend on the measurement of spatial variation in resource patterning over large areas, something that is extremely difficult to do. Ecologists measure such variation by counting and weighing types and numbers of plants, an expensive and time-consuming process even in small test plots. In addition, there is the very real danger in on-the-ground efforts of becoming "too close" to the data, of placing emphasis on taxonomy and the specific properties of individual taxa as "determinants," to the detriment of a wider perspective. For both economy of effort and maintenance of a regional perspective, remote sensing methods may be superior to on-the-ground ecological measurements of environmental diversity.

Remote sensor imagery, particularly photographic or multispectral representations of ground scenes, contains information on the reflectivity of different parts of a scene covering a portion of the earth's surface. Reflectivity is determined by ground cover, soil type, topography, and an amalgam of other natural factors—all of which would correspond to a greater or lesser extent with the distribution of vegetation. Since animal life is dependent upon the patterning of primary producers, remote sensor data should convey information about faunal resource distributions as well.

The limits of 10–100 km suggested by Harpending and Davis (1977) as a relevant distance for the discussion of resource periodicities among human groups cover a significantly larger span than do most aircraft platform images. For this reason, Landsat or other satellite scanner data may be the ideal media for experiments in the measurement of archaeologically relevant environmental diversity. One objection often raised concerning Landsat MSS data is its low resolution, so a consideration of the sufficiency of these data for spectral analysis of the sort discussed above is perhaps in order.

As will be discussed later in this chapter, the periodicities of occurrence of resources or of the landform characteristics that determine the distribution of resources constitute one property of the environment that can be measured to arrive at data that qualify as ecosystemic. For instance, the ecosystemic properties of an area may be very different if there are five apple trees and five orange trees than if there are 500 orange trees and 500 apple trees. A rule of thumb for the measurement of periodicities from serial data, the Nyquist criterion (Gillespie 1980:149), holds that at least two samples per cycle of the highest spatial frequency information to be obtained from an image are required. A Landsat MSS image provides a ground coverage of approximately 185 by 185 km; to detect a 10 km spatial period, then, $(2 \times 18)^2$ or 1369 samples would have to be derived from the

frame. Landsat MSS imagery contains some 1.6×10^7 pixels per frame, nearly 8000 times as many potential samples as would be required for such sampling. Data derived through aerial photography are even more detailed. Conventional aerial photos contain about 4×10^8 pixels per frame, and high-resolution images have several times that many pixels (Reeves 1975:1104).

An early remote sensing experiment carried out to assess the possibility of measuring archaeologically relevant environmental diversity using aerial photographs focused on the lower Chaco River drainage and surrounding badlands and mesatop areas in northwestern New Mexico during a cultural resources survey of coal mining lands (Reher 1977). An initial hypothesis advanced as part of the explanation of Archaic site densities in the study area was that Archaic site densities should increase as a function of increasing diversity in vegetation (Reher and Witter 1977:114). This hypothesis was based on the assumption that Archaic peoples pursued a generalist subsistence strategy, relying on a wide variety of vegetal resources throughout the year. This assumption may not be totally valid or realistic, based on subsequent research (Hogan and Winter 1983; Moore and Winter 1980), but a discussion of the way in which diversity measures were obtained should help to point the way for future efforts in this direction.

Two separate data sources were used to measure vegetation diversity: on-the-ground botanical survey and the analysis of aerial photographs. The aerial photographic measurements employed 1:6000 and 1:12,000 black-and-white and color transparency aerial photos of the study area, which were analyzed using an International Imaging Systems analog image analysis system. One of the capabilities of this sysrem is a graphic readout of density changes in the emulsion of a photograph placed on a light table and viewed with a high-resolution video camera. Such a graphic readout of densities of course corresponds to differences in vegetation, topography (shadow), soils, and other proxies of environmental diversity. Each photograph from the areal coverage of the study area was placed on the light table in turn, and the density graph of a north-south line across its center was examined. Peaks in this graph with an amplitude greater than an arbitrary cutoff value were counted, thus providing a simple, efficiently derived measure of the amount of variation in density across each photographic frame. The number of such graph peaks counted was assigned as a "diversity index" to the area covered by each photo frame (Ebert and Hitchcock 1977:212).

A vegetative diversity index was independently derived from analysis of plant communities and associations measured on the ground; this index was found to correspond closely with the remote-sensing-derived index. Correlation of both indices with Archaic site location data derived through transect survey indicated that Archaic site density was highest in areas lying immediately adjacent to high vegetation or environmental diversity areas, but that the sites were not necessarily within these areas themselves. A possible explanation is that high-diversity areas are extremely variable topographically and have active erosional and aggradational regimes. Thus, such areas may be inappropriate places to locate residential sites, or the archaeological record in such areas may be obliterated or hidden (Reher and Witter 1977).

In 1979, a cooperative study to further investigate the use of remote sensor data, this time from Landsat MSS, for measuring environmental diversity for archaeological purposes was initiated by the National Park Service's Branch of Remote Sensing and the U.S. Geological Survey's EROS Program. It was proposed that this study would incorporate analysis of five 500 by 500 pixel (approximately 27.5 by 27.5 km) Landsat 3 MSS subscenes in the San Juan Basin near the 1977 Chaco River study area described above. The derivation of a diversity measure from these subscenes was to be digital, and the diversity measure so derived was to be compared with an extensive archaeological computer data base that had recently been made available by the Park Service's Southwestern Regional Office in Santa Fe.

Digital analysis was undertaken at the EROS Data Center, a U.S. Geological Survey facility in Sioux Falls, South Dakota, using two digital image analysis systems, the General Electric Image 100 system and the ESL IDIMS (Interactive Digital Image Manipulation System). Subscenes were extracted from a Landsat 3 MSS tape (data collected August 3, 1979) and rerecorded onto digital tape. These data were then analyzed using a maximum likelihood cluster classifier on the IDIMS system. A 50 by 80 pixel area from each subscene that was judged to be representative of the variation within that subscene was first selected by the operators based on the ecologic/cover-type classification of the San Juan Basin discussed above (Camilli 1984). This small area was then randomly sampled to derive a training set of 5 percent, or 20 by 4 pixels. A total of 164 such samples were derived from the four subscenes. Using these samples as training sets, an unsupervised classification was performed, and 13 classes resulted. These classes were interpreted and collapsed by the operators, again on the basis of the previous cover-type interpretation as well as internalized knowledge of the area, into seven new cover types, which were then mapped as zones (Figure 9.13).

Once these steps had been completed, the EROS Data Center's Burroughs 7600 computer was used to pass a 3 by 3 pixel filter through the seven-zone classified image. For each nine-pixel area, the central pixel was replaced with a value of 0–6, indicating the number of classes other than the class represented by the central pixel that were found within the filter. This resulted in the generation of a diversity index (Figure 9.14), but unfortunately, edge effects relating to the direction that the pixel passed through the data set were introduced into the results. Attempts were made to correct for this, but the configuration of the computer system at that time was such that it could not be adapted to solve the problems. For this reason the proposed correlations between site occurrence and the diversity measure were never completed, although the method itself shows considerable promise.

A number of things can be said about and learned from this last attempt at measuring environmental diversity as an ecosystemic variable with archaeological relevance. The first is that problems of coordination and equipment compatibility sometimes make it simpler and more cost-effective for a manager to contract with an accountable scientist from the private sector for remote sensing research than to rely on cooperative, interagency agreements.

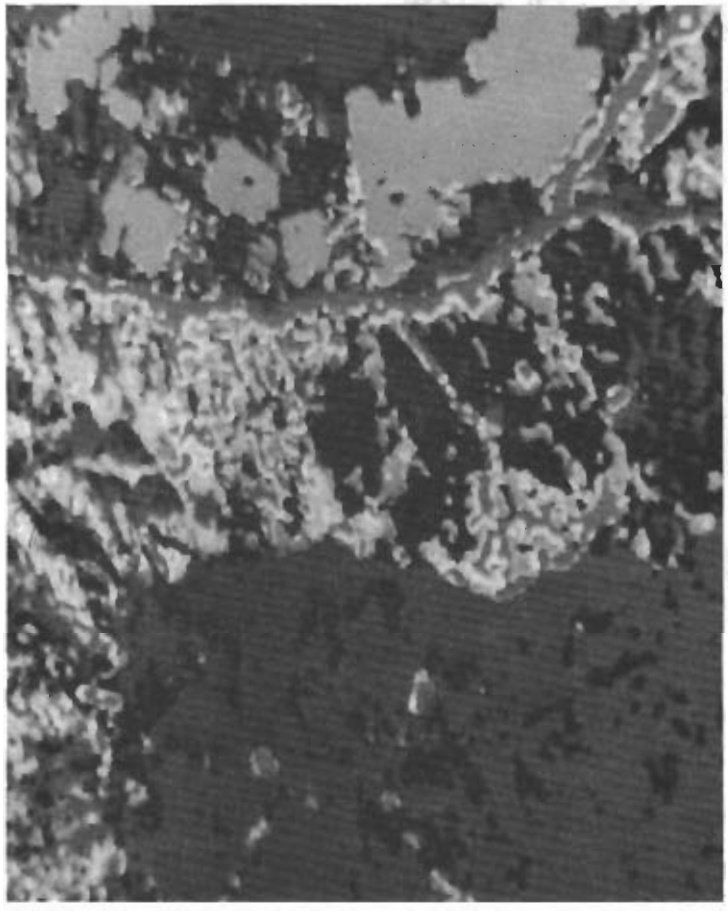


Figure 9.13. A computer-generated image showing a digital classification of cover-type units along Gallegos Wash in the San Juan Basin, northwestern New Mexico. Landsat 3 MSS data were analyzed using the IDIMS computer system at the EROS Data Center in Sioux Falls, South Dakota. The results of an attempt to derive ecosystem variables from this classification are discussed in the text.

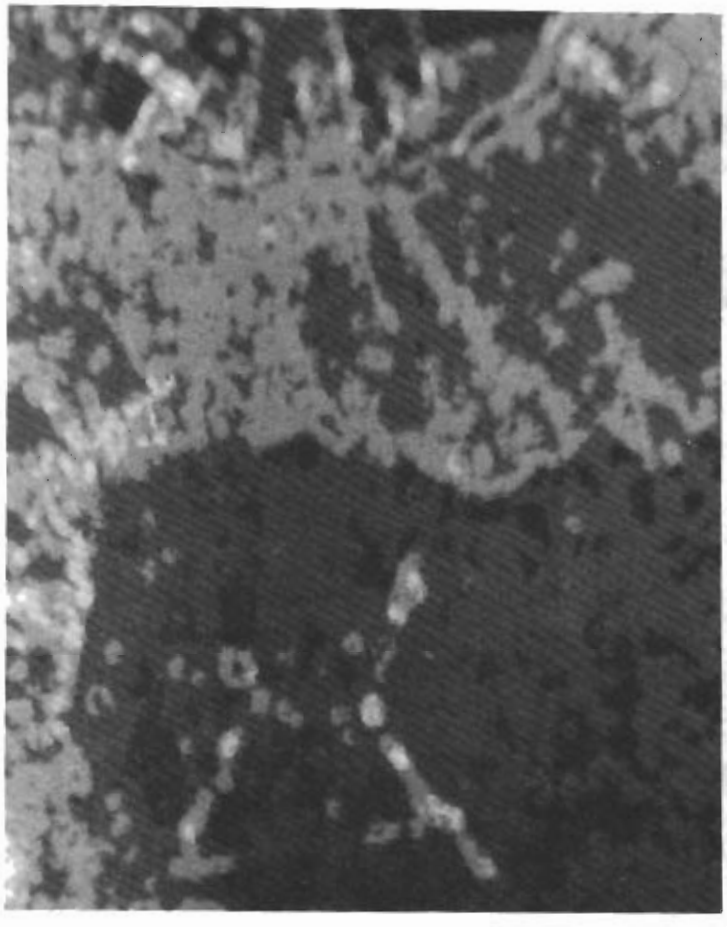


Figure 9.14. A digitally derived environmental diversity index resulting from further computer analysis of the cover-type classification shown in Figure 9.13. For each pixel in the classified scene the number of cover types occurring within a three-pixel radius was counted; this score was used to derive a diversity index. The darkest areas have the lowest diversity and the lightest areas have the highest diversity. Much past systemic behavior, including site location choice, may be more attributable to ecosystemic variables, such as diversity, than to specific vegetation or other resource composition, as discussed in the text.

The second observation that might be made is that the technology of digital analysis of remote sensor data is changing so rapidly as to make analyses that were not practical using million-dollar systems only a few years ago possible today on small, stand-alone image processors. The RIPS (Remote Image Processing System) that Charles Robinove (1986) used to derive his Landsat-based diversity index in 1984 is now available to the general public as a $5000 add-on to most personal computers. This diversity measurement attempt also illustrates at least one application of remote sensing in which digital, pixel-by-pixel classification of data is far more useful than visual interpretation of an image into zones or areas of assumed significance, for it would be impossible to pass a filter through an image if it were not composed of pixels.

Finally, this example emphasizes the fact that remote-sensing-based approaches to the measurement of ecosystem variables for prediction and modeling have not been perfected, and that it may not be easy to perfect them. Remote sensing approaches, like predictive modeling in general, can only be refined through cooperative research and development on the part of managers and archaeologists.

The last point is one in which remote sensing can, I feel, play an especially important role in uniting the efforts of managers and archaeologists. Remote sensor data forms an integral and all-important part of most geographic information systems (as discussed in Chapter 10). Such systems have been undergoing intensive development, particularly by natural resource managers and scientists, for at least a decade. I see focus on remote sensing as a primary data source for predictive experiments in archaeology as one way of developing a common ground, an independent data base, and ultimately an analytical tool that can be shared by archaeologists, natural resource scientists, and managers. Such a common interest could do much toward uniting cultural and natural resource management.

REFERENCES CITED

Avery, T. Eugene
1977 *Interpretation of Aerial Photographs*, 3rd rev. ed. Burgess, Minneapolis.

Avery, T. Eugene, and Thomas R. Lyons
1981 *Remote Sensing: Aerial and Terrestrial Photography for Archeologists.* Cultural Resource Management Division, National Park Service, Washington, D.C.

Baker, Craig, and Steven E. Sessions
1979 The Environmental Context of Archaeological Sites in the EPCC Survey Area. In *The Archaeology of Southwest Gallegos Mesa: The EPCC Survey*, edited by Steven E. Sessions. Navajo Nation Papers in Anthropology 1. Window Rock, Arizona.

1982 Predictive Model Development and Evaluation. In *Archaeological Variability Within the Bisti-Star Lake Region, Northwestern New Mexico*, edited by Meade F. Kemrer, pp. 121–173. ESCA-Tech Corporation, Albuquerque.

Beazeley, G. A.
1919 Air Photography in Archaeology. *Geographical Journal* 53:330–335.

Bernstein, Ralph
1983 Image Geometry and Rectification. In *Manual of Remote Sensing*, 2nd ed., edited by Robert N. Colwell, pp. 873–992. American Society of Photogrammetry, Falls Church, Virginia.

Billingsley, Fred C.
1983 Data Processing and Reprocessing. In *Manual of Remote Sensing*, 2nd ed., edited by Robert N. Colwell, pp. 719–792. American Society of Photogrammetry, Falls Church, Virginia.

Binford, Lewis R.
1982 The Archaeology of Place. *Journal of Anthropological Archaeology* 1:5–31.

Brown, G. N., and J. I. Ebert
1978 Ecological Mapping for Purposes of Sample Stratification in Large-Scale Cultural Resources Assessment: The National Petroleum Reserve in Alaska. In *Remote Sensing and Nondestructive Archeology*, edited by Thomas R. Lyons and James I. Ebert, pp. 53–63. National Park Service, Washington, D.C.

Butzer, K. W.
1977 Geo-archaeology in Practice. *Reviews in Archaeology* 4:125–131.

Camilli, Eileen
1979a Remote Sensing and Cultural Resources Management for the Navajo Indian Irrigation Project (NIIP). In *Report of the Cultural Resources Survey of Blocks VI and VII of the NIIP*, by William Reynolds, Eileen Camilli, and David Mayfield. ESCA-Tech Corporation, Albuquerque, New Mexico, and Costa Mesa, California.

1979b The Sampling Design and Known Cultural Resource Locations. In *Preliminary Report: Reconnaissance Survey of the Archaeological and Related Resources of the Lower Puerco and Salado Drainages, Central New Mexico*, edited by Mark Wimberly and Peter Eidenbach, pp. 24–34. Human Systems Research, Tularosa, New Mexico.

1984 An Ecological Cover-Type Map of the San Juan Basin, Northwestern New Mexico. In *Remote Sensing in Cultural Resource Management: The San Juan Basin Project*, edited by Dwight L. Drager and Thomas R. Lyons, pp. 39–56. Cultural Resource Management Division, National Park Service, Washington, D.C.

Camilli, Eileen, and Tim Seaman
1979 Preliminary Assessment of Cultural Resources for the Crownpoint to Bisti Water Pipeline Route: Cultural Resources Management with the Aid of Remote Sensing. In *New Mexico Generating Station Environmental Assessment*. Public Service Company of New Mexico, Albuquerque.

Capper, J. C.
1907 Photographs of Stonehenge, as Seen from a War Balloon. *Archaeologia* 60:571, Plates 69–70.

Carstens, Kenneth C., Thomas C. Kind, and Neil V. Weber
1981 Using Remote Sensing in a Predictive Archaeological Model: The Jackson Purchase Region, Kentucky. *Proceedings of the Pecora VII Symposium, Sioux Falls, South Dakota, October 18–21, 1981*, pp. 494–507. American Society of Photogrammetry, Falls Church, Virginia.

Colwell, Robert N., editor
1983 *Manual of Remote Sensing*, 2nd ed., 2 vols. American Society of Photogrammetry, Falls Church, Virginia.

Conkey, Margaret W.
1980 The Identification of Prehistoric Hunter-Gatherer Aggregation Sites: The Case of Altamira. *Current Anthropology* 21:609–620.

Cook, J. P., and W. J. Stringer
1975 *Feasibility Study for Locating Archaeological Village Sites by Satellite Remote Sensing.* Geophysical Institute, University of Alaska, Fairbanks. Contract NAS5-21833, ERTS Project 110-N. Submitted to NASA Goddard Space Flight Center, Greenbelt, Maryland.Copies available at the Geophysical Institute, University of Alaska, Fairbanks.

Custer, Jay F., and George J. Galasso
1983 A Prehistoric Archaeological Survey of the St. Jones and Murderkill Drainages, Kent County, Delaware. *Bulletin of the Archaeological Society of Delaware* n.s. 14:1-18.

Custer, Jay F., Timothy Eveleigh, and Vyatautas Klemas
1983 A Landsat-Generated Predictive Model for Prehistoric Archaeological Sites in Delaware's Coastal Plain. *Bulletin of the Archaeological Society of Delaware* n.s. 14:19-38.

Custer, Jay F., Patricia Jehle, Thomas Klatka, and Timothy Eveleigh
1984 A Cultural Resources Reconnaissance Planning Srudy of the Proposed Route 13 Relief Corridor, New Castle and Kent Counties, Delaware. Submitted to the Location and Environmental Studies Office, Division of Highways, Delaware Department of Transportation, Dover. Copies available at the Location and Environmental Studies Office, Division of Highways, Delaware Department of Transporration, Dover.

Dowman, I. J.
1980 Review of T. R. Lyons and T. E. Avery, *Remote Sensing: A Handbook for Archeologists and Cultural Resource Managers* and *Supplement No. 1: Remote Sensing. Practical Exercises in Remote Sensing in Archeology. Aerial Archaeology* 4:95.

1983 Review of T. E. Avery and T. R. Lyons, *Supplement No. 7: Remote Sensing. Aerial and Terrestrial Photography for Archaeologists. Aerial Archaeology* 7:26.

Drager, Dwight L.
1980a Projecting Archeological Site Concentrations from Cover-Type Maps Developed from Remote Sensing Data. In *Proceedings of the Second Conference on Scientific Research in the National Parks Held at San Francisco, California, on November 26-30, 1979,* vol. 1, pp. 151-157. National Park Service, Washington, D.C.

1980b Vegetative Stratifications from Aerial Imagery. In *Cultural Resources Remote Sensing,* edited by T. R. Lyons and F. J. Marhien, pp. 79-86. National Park Service, Washington, D.C.

Drager, Dwight L., and Arthur K. Ireland
1986 *The Seedskadee Project: Remote Sensing in Non-Site Archeology.* National Park Service, Albuquerque, and Bureau of Reclamation, Salt Lake City.

Drager, Dwight L., and Thomas R. Lyons, editors
1983 *Remote Sensing in Cultural Resource Management: The San Juan Basin Project.* Cultural Resources Management Division, National Park Service, Washington, D.C.

Drager, Dwight L., James I. Ebert, and Thomas R. Lyons
1982 Remote Sensing and Nondestructive Archeology: Approaches to Cultural Resource Management. In *The San Juan Basin Tomorrow: Planning for the Conservation of Cultural Resources in the San Juan Basin,* edited by Fred Plog and Walter Wait, pp. 219-244. National Park Service, Southwest Region, Albuquerque, in cooperation with the School of American Research, Santa Fe.

Dunnell, Robert C.
1980 Remote Sensing in Archaeological Sampling Designs. *Transportation Engineering Journal* 106(TE3):349-363. American Society of Civil Engineers Proceedings Paper 15248.

Ebert, James I.
1978 Remote Sensing and Large-Scale Cultural Resources Management. In *Remote Sensing and Nondestructive Archeology,* edited by Thomas R. Lyons and James I. Ebert, pp. 21-34. National Park Service, Washington, D.C.

1980 Remote Sensing in Large-Scale Cultural Resources Survey: A Case Study from the Arctic. In *Cultural Resources Remote Sensing*, edited by T. R. Lyons and F. J. Mathien, pp. 7–54. National Park Service, Washington, D.C.

1984 Remote Sensing Applications in Archaeology. In *Advances in Archaeological Method and Theory*, vol. 7, edited by Michael B. Schiffer, pp. 293–362. Academic Press, New York.

Ebert, James I., and Galen N. Brown

1981 Remote Sensing in the NPR-A Cultural Resources Assessment. In *Slogging, Humping, and Mucking Through the NPR-A*, edited by Craig W. Davis, Dana C. Linck, Kenneth M. Schoenberg, and Harvey M. Schields, pp. 349–419. Anthropology and Historic Preservation Cooperative Park Studies Unit, University of Alaska, Fairbanks.

Ebert, James I., and Alberto A. Gutierrez

1979a An Approach to the Application of Remote Sensor Data to Cultural Resources Sampling in Shenandoah National Park. Ms. on file, Division of Remote Sensing, Southwest Cultural Resources Center, University of New Mexico, Albuquerque.

1979b Relationships between Landscapes and Archaeological Sites in Shenandoah National Park: A Remote Sensing Approach. *APT: Bulletin of the Association for Preservation Technology* 11(4):69–87.

1979c The Application of Remote Sensor Data to Cultural Resources Sampling and Prediction in Shenandoah National Park. Ms. on file, Division of Remote Sensing, Sourhwest Cultural Resources Center, University of New Mexico, Albuquerque.

1980 Cultural Resources and Remote Sensing in the Eastern Deciduous Woodland: Experiments at Shenandoah National Park. Paper presented at the Second Conference on Scientific Research in the National Parks, Washington, D.C.

1981 Remote Sensing of Geomorphological Factors Affecting the Visibility of Archaeological Materials. In *Technical Papers of the American Society of Photogrammetry 47th Annual Meeting, Washington, D.C., February 22–27, 1981*, pp. 226–236. American Society of Photogrammetry, Falls Church, Virginia.

Ebert, James I., and Robert K. Hitchcock

1977 The Role of Remote Sensing. In *Settlement and Subsistence in the Lower Chaco River: The CGP Survey*, edited by Charles A. Reher, pp. 191–216. University of New Mexico Press, Albuquerque.

Ebert, James I., and Thomas R. Lyons, editors

1983 Archaeology, Anthropology, and Cultural Resources Management. In *Manual of Remote Sensing*, 2nd ed., vol. 2, edited by Robert N. Colwell, pp. 1233–1304. American Society of Photogrammetry, Falls Church, Virginia.

Ehrenhard, Ellen B.

1980 Remote Sensing and the Predictive Site Location Model (for Management), Chattahoochee River National Recreation Area. Ms. on file, Southeast Archeological Center, National Park Service, Tallahassee.

Ehrenhard, John E.

1980 Cultural Resource Inventory of the Big Cypress Swamp: An Experiment in Remote Sensing. In *Cultural Resources Remote Sensing*, edited by T. R. Lyons and F. J. Mathien, pp. 105–117. National Park Service, Washington, D.C.

Elston, Robert G., David S. Whitley, Michael S. Berry, Alan S. Lichty, Michael P. Drews, and Charles D. Zeier

1983 Class II Archaeological Survey of Selected Portions of the Naval Weapons Center, China Lake, California. Intermountain Research, Silver City, Nevada. Prepared for Michael Brandman Associates, Irvine, California.

Estes, John E., Earl J. Hajic, and Larry R. Tinney
1983 Fundamentals of Image Analysis: Analysis of Visible and Thermal Infrared Data. In *Manual of Remote Sensing*, 2nd ed., vol. 2, edited by Robert N. Colwell, pp. 987–1124. American Society of Photogrammetry, Falls Church, Virginia.

Evans, R.
1983a Review of S. A. Morain and T. K. Budge, *Supplement No. 2: Remote Sensing. Instrumentation for Nondestructive Exploration of Cultural Resources. Aerial Archaeology* 7:24.

1983b Review of T. R. Lyons, R. K. Hitchcock, and W. H. Wills, *Supplement No. 3: Aerial Anthropological Perspectives. A Bibliography of Remote Sensing in Cultural Resource Studies. Aerial Archaeology* 7:24.

Findlow, Frank J., and Linda Confeld
1980 Landsat Imagery and the Analysis of Archaeological Catchment Territories: A Test of Method. In *Catchment Analysis: Essays on Prehistoric Resource Space*, edited by Frank J. Findlow and Jonathon E. Ericson, pp. 31–52. Anthropology UCLA 10(1–2).

Foley, Robert
1981 Off-Site Archaeology: An Alternative Approach for the Short-Sited. In *Patterns of the Past: Essays in Honor of David L. Clarke*, edited by I. Hodder, G.Ll. Isaac, and N. Hammond, pp. 157–183. Cambridge University Press, Cambridge, England.

Fuller, R. M.
1983 Review of S. A. Morain, T. K. Budge, and A. Komarek, *Supplement No. 5: Remote Sensing. An Airborne Spectral Analysis of Settlement Sites in the Chaco Canyon and Bandelier National Monuments. Aerial Archaeology* 7:25.

Gal, Robert
1979 Archaeological Implications of Small-Scale False-Color Infrared Photography in the National Petroleum Reserve-Alaska. Paper presented at the 6th Annual Meeting of the Alaska Anthropological Association, April 6, 1979, Fairbanks.

Gillespie, Alan R.
1980 Digital Techniques of Image Enhancement. In *Remote Sensing in Geology*, edited by Barry S. Siegel and Alan R. Gillespie, pp. 139–226. Wiley, New York.

Gladfelter, Bruce G.
1981 Developments and Directions in Geoarchaeology. In *Advances in Archaeological Method and Theory*, vol. 4, edited by Michael B. Schiffer, pp. 343–364. Academic Press, New York.

Haase, William
1981 Estimates of Archaeological Site Density and Distribution on Cedar Mesa, Utah: A Study Based on Photomosaic Interpretation and Probability Sampling. Ms. on file, Department of Anthropology, Washington State University, Pullman.

Haralick, Robert M., and King-Sun Fu
1983 Pattern Recognition and Classification. In *Manual of Remote Sensing*, 2nd ed., vol. 1, edited by Robert N. Colwell, pp. 793–805. American Society of Photogrammetry, Falls Church, Virginia.

Harpending, Henry, and Herbert Davis
1977 Some Implications for Hunter-Gatherer Ecology Derived from the Spatial Structure of Resources. *World Archaeology* 8:275–286.

Hogan, Patrick, and Joseph C. Winter, editors
1983 *Economy and Interaction along the Lower Chaco River: The Navajo Mine Archeological Program.* Office of Contract Archeology and Maxwell Museum of Anthropology, University of New Mexico, Albuquerque.

Holmer, Richard N., editor
1982 *Castle Valley Rail Line, Denver and Rio Grande Western Railroad: Known Archaeological and Historical Site Evaluation and Predictive Model.* University of Utah Archaeological Center. Submitted to Sverdrup & Parcel and Associates, St. Louis. Copies available at University of Utah Archaeological Center, Provo.

Inglis, Michael, Thomas K. Budge, and James I. Ebert
1984 Preliminary Evaluation of Simulated SPOT Data for Cultural Resources Assessment: An Experiment at Bandelier National Monument, New Mexico. Paper presented at the 1984 SPOT Symposium, May 20-23, Scottsdale, Arizona. Copies available from Technology Applications Center, University of New Mexico.

Kemrer, Meade F., editor
1982 *Archaeological Variability within the Bisti-Star Lake Region, Northwestern New Mexico.* ESCA-Tech Corporation. Submitted to the Bureau of Land Management, Albuquerque District. Copies available at Albuquerque District Office, Bureau of Land Management.

Kruckman, Laurence
1972 The Techniques and Application of Aerial Photography to Anthropology: A Bibliography. *Exchange Bibiliography* 339. Council of Planning Librarians, Monticello, Illinois.

Lade, K.-Peter
1981a Surveying by Satellite: A Preliminary Report on the Utility of rhe Multispectral Scanner Aboard Landsat to Archaeological Surveying in Maryland. Paper presented at the Middle Atlantic Archaeological Conference, March 20-22, Ocean City, Maryland.

1981b Surveying by Satellite in Maryland. Richard B. Stearns Memorial Lecture delivered at the 16th Annual Spring Symposium, Archaeological Society of Maryland, April 25, 1981.

1982 The Monitoring and Detection of Archaeological Resources in Maryland Through the Use of Landsat Digitized Multispectral Data. Paper presented at the 47th Annual Meeting of the Society for American Archaeology, Minneapolis.

Lillesand, Thomas M., and Ralph W. Kiefer
1979 *Remote Sensing and Image Interpretation.* John Wiley and Sons, New York.

Lindbergh, C. A.
1929 The Discovery of Ruined Maya Cities. *Science* 70:12-13.

Lyons, T. R., and J. I. Ebert, editors
1978 *Remote Sensing and Nondestructive Archeology.* Cultural Resources Management Division, National Park Service, Washington, D.C.

Madry, Scott L. H.
1983 *A Multiscalar Application of Remote Sensing in a Temperate Regional Archaeological Survey.* M.A. Thesis, Department of Anthropology, University of North Carolina, Chapel Hill.

1984 A Multiscalar Application of Remote Sensing in a Temperate Regional Archaeological Survey. *Technical Papers of the 50th Annual Meeting of the American Society of Photogrammetry, March 11-16, 1984, Washington, D.C.*, pp. 113-121.

May, John R.
1978 Appendix A: Sources of Available Remote Sensor Imagery, and Appendix B: Sources of New Imagery Missions. In *Guidance for Application of Remote Sensing to Environmental Management.* Instruction Report M-78-2. U.S. Army Waterways Experiment Station, Vicksburg, Mississippi.

McAnany, Patricia A., and Ben A. Nelson
1982 *Cultural Resources Overview for the New Mexico Synfuels Project.* Office of Contract Archeology, University of New Mexico, Albuquerque.

Moore, James L., and Joseph C. Winter, editors
1980 *Human Adaptations along the Lower Chaco River: The UII Mitigation Project.* Office of Contract Archeology, University of New Mexico, Albuquerque.

Parrington, Michael
1983 Remote Sensing. *Annual Review of Anthropology* 12:105–124.

Perisset, M. C., and A. Tabbagh
1981 Interpretation of Thermal Prospection on Bare Soils. *Archaeometry* 23:169–188.

Plog, Fred, and Walter Wait, editors
1982 *The San Juan Basin Tomorrow: Planning for the Conservation of Cultural Resources in the San Juan Basin.* National Park Service, Southwest Region, Albuquerque, in cooperation with the School of American Research, Santa Fe.

Ray, Richard G.
1980 *Aerial Photographs in Geologic Interpretation and Mapping.* U.S. Geological Survey Professional Paper 373.

Reeves, Robert G., editor
1975 *Manual of Remote Sensing,* 1st ed. American Society of Photogrammetry, Falls Church, Virginia.

Reher, Charles A.
1977 *Settlement and Subsistence in the Lower Chaco River: The CGP Survey.* University of New Mexico Press, Albuquerque.

Reher, Charles A., and Daniel C. Witter
1977 Archaic Settlement and Vegetative Diversity. In *Settlement and Subsistence in the Lower Chaco River: The CGP Survey,* edited by Charles A. Reher, pp. 113–126. University of New Mexico Press, Albuquerque.

Riley, Derrick N.
1980 *Early Landscape from the Air.* Sheffield, England.

1982 *Aerial Archaeology in Britain.* Shire Publications, Aylesbury, England.

Robinove, Charles J.
1986 Spatial Diversity Index Mapping of Classes in Grid Cell Maps. *Photogrammetric Engineering and Remote Sensing* 52:1171–1173.

Schiffer, Michael B.
1983 Toward the Identification of Formation Processes. *American Antiquity* 48:675–706.

Tabbagh, A.
1977 *Sur la determination du moment de mesure favorable et l'interpretation des resultats en prospection thermique archeologique.* Annales de Geophysique 33:179–188.

Tandarich, John Paul
1975 *Correlation of Soil Types and Archaeological Site Location in Webster County, Iowa.* M.S. thesis, Department of Anthropology, University of Iowa, Ames.

Vogt, E. Z.
1974 *Aerial Photography in Anthropological Field Research.* Harvard University Press, Cambridge.

Wandsnider, LuAnn, and James I. Ebert
1983 Modeling Climatic and Landform Factors Affecting the Character of the Archaeological Record in Arid Lands: A Remote Sensing Approach in Southwestern Wyoming. Paper presented at the Silver Anniversary Meetings of the Western Social Science Association and the Association for Arid Lands Studies, April 27–30, Albuquerque.

Wells, Ian

1981 *A Spatial Analysis Methodology for Predicting Archaeological Sites in Delaware and its Potential Application in Remote Sensing.* M.A. thesis, Department of Maritime Studies, University of Delaware, Wilmington.

Wells, Ian, Jay Custer, and Vic Klemas

1981 Locating Prehistoric Archaeological Sites Using Landsat. Paper presented at the Fiftieth International Symposium on Remote Sensing of the Environment, Ann Arbor.

Whimster, R. P.

1983a Review of C. Baker and G. J. Gumerman, *Supplement No. 6: Remote Sensing. Archaeological Applications of Remote Sensing in the North Central Lowlands. Aerial Archaeology* 7:25–26.

1983b Review of C. M. Aikens, W. G. Loy, M. D. Southard, and R. C. Haines, *Supplement No. 4: Remote Sensing. Oregon. Aerial Archaeology* 7:24.

Wilson, D. R.

1982 *Air Photo Interpretation for Archaeologists.* St. Martin's Press, New York.

Wood, W. R., and D. L. Johnson

1978 A Survey of Disturbance Processes in Archaeological Site Formation. In *Advances in Archaeological Method and Theory,* vol. 1, edited by Michael B. Schiffer, pp. 315–381. Academic Press, New York.

Chapter 10

GEOGRAPHIC INFORMATION SYSTEMS: TECHNICAL AIDS FOR DATA COLLECTION, ANALYSIS, AND DISPLAY

Kenneth L. Kvamme and Timothy A. Kohler

INTRODUCTION

Timothy A. Kohler

By this time we have seen that predictive modeling of archaeological resources may involve consideration of the characteristics of catchments around potential site locations, of distances to various resource types from potential locations, and of various characteristics of the potential site location itself. Maps of several different resource types and landscape characteristics may each need to be analyzed in terms of catchment, distance, and point characteristics. Locations satisfying certain criteria on all these maps may need to be identified and located. Geographic information systems are a computerized aid for the collection, management, analysis, and display of the large sets of spatially referenced data required for such projects. This chapter begins with an overview of what these systems can do and then explains their various capabilities in more detail.

Beyond its obvious role in helping to organize, overlay, and display data, a geographic information system (GIS) also may help agencies to make cultural resource management survey and predictive modeling efforts both more comparable from project to project and more cumulative in their results. At present, the physical models—maps—produced by various archaeological consultants are drawn to different scales, using different standards. If instead models were based on a single GIS, or on compatible systems at identical resolutions, then they could be readily compared, and the predictions made by one group of modelers could be tested by later surveys more accurately and conveniently. Moreover, models could be easily refined and remapped, and the results of these refinements (and the differences between versions) would be readily apparent. A good case could be made that either agencies should maintain their own GIS and require all contractors to work on it, or they should maintain long-term arrangements with contractors for the construction of data bases containing environmental data, site location information, and predictive models so that the cycle of model construction, testing,

revision, and verification could be carried forward cumulatively. This chapter, however, considers only the technical role, not the implications for policy, that geographic information systems may have in the predictive modeling process.

Maps can be defined as scales for measuring the property of *location* for some attribute (Lewis 1977:3-10). Map data differ from other data in that the location of each feature relative to all others is maintained, making properties of location (such as distance) readily available for study. Most of the large computerized software systems that archaeologists use regularly (such as SPSS and SAS) ordinarily maintain information in a sequentially organized data base. Location can be entered in such a data base by introducing variables for northing and easting, for example, but the internal organization of the data base usually remains random with respect to these variables, and analysis of locational properties is cumbersome.

In a GIS, on the other hand, the internal organization of the data either mimics that of the map from which it is distilled or is based on other conventions that allow the spatial structure of the mapped attribute to be easily reconstructed. This facilitates various spatial studies, such as those requiring distance measures (including catchment studies), and permits overlaying of various maps on top of each other so that the spatial interaction of the mapped attributes can be studied.

A working GIS consists of software (computer programs), the hardware on which that software operates, and a spatial data base, but the term *GIS* is often used to refer only to the software used for data entry, management, manipulation, analysis, and display. Many geographic information systems have separate systems, or subprograms, for these various major functional categories. There are probably well more than 100 geographic information systems in use around the world, in many times that number of installations; access to a GIS by researchers and managers in university and agency contexts will soon be commonplace.

Comparative reviews of the most common systems are now available: Hansen (1983) compares MOSS/MAPS with IDIMS; several systems that were originally designed to process remote sensing data, including VICAR and IDIMS, are compared by Bracken et al. (1983); and Erikson et al. (1983) discuss three microcomputer-based geographic information systems. Munro (1983) draws on the experience of a large corporation in suggesting how a suitable GIS can be objectively selected from those available. Systems used by the Dominion of Canada and by the states of New York and Minnesota are described by Tomlinson et al. (1976). Finally, the American Farmland Trust (1985) tabulated costs, operating environments, and data entry, editing, updating, retrieval, analysis, output, and display functions for 65 geographic information systems, include 16 operating on microcomputers. Even such a recent publication is already somewhat out-of-date, however, as both hardware and software developments in this field are occurring very rapidly.

Training in the structure and use of geographic information systems is available from several sources (Table 10.1). Articles relevant to geographic information systems appear regularly in the journals and conference proceedings listed in Table 10.2, and Estes et al. (1984) and Marble et al. (1984) present useful collections of GIS-related articles.

TABLE 10.1.

Selected training opportunities in geographic information systems

Organization	*System (if any)*
Training and Assistance Office U.S. Geological Survey EROS Data Center Sioux Falls SD 57198 (605) 594-6114	IDIMS
Remote Sensing Institute South Dakota State University P.O. Box 507 Brookings SD 57007 (605) 688-4814	AREAS
Yale University School of Forestry and Environmental Studies 205 Prospect St. New Haven CT 06511 (203) 436-0440	MAP
Laboratory for Application of Remote Sensing Data Purdue University 1291 Cumberland Ave. West Lafayette IN 47906 (317) 494-6305	LARSYS
Continuing Engineering Education Program George Washington University Washington, D.C. 20052 (202) 676-6106	
Graphics and Image Analysis Group Computing Service Center Washington State University Pullman WA 99164-1220 (505) 335-0411	VICAR/IBIS
U.S. Fish and Wildlife Service Division of Biological Services Western Energy and Land Use Team Drake Creekside One, 2627 Redwing Rd. Ft. Collins CO 80526-2899	MOSS/MAPS
Geographic Information Systems Laboratory Central Washington University Ellensburg, WA 98926 (509) 963-1914	GRASS

TABLE 10.2

Selected journals and conference proceedings containing more advanced discussions of geographic information systems

JOURNALS

Area
Canadian Cartographer
Computer Vision Graphics and Image Processing
Computers and Geosciences
Computers, Environment, and Urban Systems
Environment
Environmental Management
Geo
Geographical Analysis
Geo-Processing
IEEE Transactions on Geoscience and Remote Sensing
IEEE Transactions on Pattern Analysis and Machine Intelligence
International Journal of Remote Sensing
Photogrammetric Engineering and Remote Sensing
Remote Sensing of Environment
Soil Survey and Land Evaluation

PAPERS AND PROCEEDINGS

International Symposium on Computer-Assisted Cartography
International Symposium on Cartography and Computing
International Symposium on Remote Sensing of the Environment
International Symposium on Spatial Data Handling
Annual Meeting of the American Society of Photogrammetry
Proceedings of the Pecora Symposium

ABSTRACTS

Geo Abstracts, G: Remote Sensing, Photogrammetry, and Cartography

THE POTENTIAL OF GEOGRAPHIC INFORMATION SYSTEMS FOR RESEARCH, DEVELOPMENT, AND APPLICATION OF ARCHAEOLOGICAL SITE LOCATION MODELS

Kenneth L. Kvamme

The Need for Geographic Information System Techniques

In the previous chapters, several methods and models for classifying a location or region as site-likely (or site-type-likely) were introduced. All of these procedures are based, at least during some stage of the modeling process, on measured data (where measurements can also refer to nominal-level class categories), and many

require large numbers of calculations. The various quantitative approaches require measurements at each site location (e.g., of various environmental phenomena) and also at locations of background environment where sites are not present, termed *nonsites*, if a control-group approach is used during initial model development (see Chapter 8). Similarly, measured data are required for all sites (and nonsites) in model testing phases. Finally, to apply most archaeological locational models across a region of study requires tremendous numbers of measurements. For example, if a model based on several environmental variables is to be applied across some region, measurements of each variable might be required every 50 m across the region for sufficient resolution in application. The problems of making vast numbers of measurements and performing an even larger number of statistical calculations constitute the greatest difficulties in the development, testing, and practical application of many archaeological modeling strategies in regional cultural resource management contexts.

For the simplest application of environmentally based models, such variables as slope, aspect, and distance to water can be measured by hand at a specific locus on a topographic map. A site location model could then be applied to the measurements (usually requiring a few calculations) in order to assess the "site likelihood" or "site favorableness" of the location. This approach can be quite useful to cultural resource managers in assessing archaeological sensitivity at, for example, proposed well pad locations.

As the size of the area to be assessed increases, however (as the number of well pads increases and as access roads to the pads are included in the project, for example), the labor-intensive hand measurement and calculation requirements rapidly become impractical. Many projects on federal lands encompass large areas; in such cases, the logical approach would be to replicate the above procedure systematically across the area under consideration, performing the measurements (and calculations) every 50 m east-west and north-south, for example. The outcome would be a wide-area "site sensitivity surface" depicting favorable or likely locations for cultural resources based on model specifications. Needless to say, performing measurements of multiple variables at some point on a map is quite tedious; replicating this process every 50 m or so, even over a small area, is incredibly time-consuming and therefore costly. In addition, once these data have been collected, the time and expense for all of the calculations required to apply most models must be considered as well.

As an illustration of the magnitude of this problem, the effort that was required to produce a "probability surface map" of site presence for a single quarter section utilizing manual techniques can be examined (Kvamme 1983a; this map is illustrated in Figure 8.1). To produce this map, six environmental predictors (slope, aspect, view angle, shelter rank, vantage distance, and distance to water) were measured by hand at 256 points evenly spaced at 50 m intervals across the quarter section for a total of 1536 measurements. Next, the probability of each location's membership in a site-present class, conditional on the measured data, was estimated by a preestablished discriminant function. The mathematical operations needed to assess one

location required roughly eight additions, three subtractions, nine multiplications, one division, and three exponentiations; for all 256 locations approximately 6144 calculations were performed! Finally, it was necessary to produce a graphic of the result for each location, which constituted a mapping of the model; this required further effort. It is clear that application of this kind of model utilizing manual techniques is impractical for any but the smallest of regions.

Manual techniques pose a number of problems in the model development and testing stages as well. Perhaps most apparent is the effective limitation of sample sizes owing to the excessive labor requirements of measurement. For example, a region might contain several hundred known sites, but it might not be possible to use all of them for model development or testing because of the difficulties of measurement. This is even more likely to be the case for nonsites if the control-group approach is used, since potential sample sizes of many thousands of nonsites can in principle be obtained from the background environment.

Perhaps a more serious effect of hand-measurement of variables is that a large amount of variation can be introduced into an analysis simply through measurement error. Significant differences can be observed between measurements taken by different people or in measurements made by the same person at different times, even for variables as easy to measure as distance to nearest stream or slope as percent grade. This factor can introduce major variation into the outcome of a model and can also affect the application of a model.

A major disadvantage of manual measurement has become apparent only with the implementation of computer-based GIS technology in archaeological locational studies. Human measurement, primarily because it is slow and time-consuming, actually limits the kinds of phenomena that might potentially be examined, or even conceived, in site location research. For example, for a given locus on a map (such as a site location), or even for several loci, it might be possible to estimate a least-effort travel distance (as opposed to a linear distance) to a nearest water source (discussed below), or it might be possible to calculate, as a relative measure of view quality, the percentage of terrain that is visible within a given area. It is not possible to do these kinds of calculations manually for many hundreds or thousands of locations (or, for example, every 50 m across a map area). In fact, since we inherently think in a "manual mode," such variables are rarely even considered. This poses a serious constraint on archaeological locational research.

Archaeologists are great gatherers of information. We collect data pertaining to where sites are found or even where individual artifacts are located. We gather information describing regions surveyed, the intensity of the survey, when the region was surveyed, and who surveyed it. We collect data about site content, the locations of features and artifacts within a site, cultural affiliation, various site components, and the amount and kinds of work performed. Various ecological data, such as environmental associations, might be recorded, as well as modern features, such as existing roads, trails, dwellings, and towns. It is important to recognize that much, perhaps most, of our data are geographically distributed; that is, they have a mappable component. A major problem is that it is often difficult to manage large

bodies of regional information and to retrieve particular information because part of the data might exist on maps while other information might be located in site forms, in project reports, in published articles, or even in museum collections. The usefulness of our great collecting efforts is thus severely compromised.

Finally, archaeologists have been working with unmanageably large, geographically distributed computer data bases, such as digital representations of remotely sensed images or digital terrain models, for a number of years (e.g., Green and Stewart 1983; Lyons and Hitchcock 1977). These unwieldy sources of information are often difficult to analyze, explore, and manipulate, and it is not easy to arrive at conclusions about them (McLeod and Jafek 1984). Various sources of data might occur at different scales, in several map projections, or might even be geometrically distorted owing to the tilted angle of a remote sensor platform, making it difficult not only to *register* one source of data to another (such that a particular point in space lines up with the same point in all the other data sources over the entire region of study) but to locate even a single point in space in all data sources. These problems are major limiting factors in the practical use of these data bases in regional archaeological investigations.

GIS technology can virtually eliminate these problem areas and limitations. First, computers can perform many thousands of measurements of potentially all variables examined in site location studies in a matter of seconds and permanently store those measurements for later use. This virtually eliminates sample-size problems for known site locations and also permits us to obtain extremely large samples of the background environment (or nonsites) for comparative studies as well. Second, such complex calculations as probability estimates can be performed quickly and in large numbers. Third, cartographic capabilities inherent in a GIS can provide maps of virtually any result quickly and at low cost. Fourth, variation in measurement is entirely eliminated: the computer produces the same result every time. Fifth, depending on the ingenuity of the user, the available software, and the software developer, the potential for creating and exploring new types of information of relevance to archaeological research and problem solving in site location studies is limitless. Last, geographic information systems provide a comprehensive system for the management of large, diverse, and unwieldy geographic data sets obtained from virtually any source, such as site files, aerial photographs, remotely sensed imagery, or conventional maps. All types of information, despite their original disparity, are referenced to a common geographic coordinate base (such as longitude and latitude or the Universal Transverse Mercator grid), providing a logical means for data storage, manipulation, retrieval, and interpretation. Thus, only through GIS capabilities does it become possible to utilize many of the data and approaches toward understanding and modeling prehistoric site distributions that have been outlined in this volume.

The following sections describe in greater detail the mechanics behind geographic information systems and their capabilities for archaeological locational research. The material in these sections is not necessarily limited to a discussion of what existing geographic information systems are able to do. Rather, the goal is to

present what a GIS can *potentially* offer to archaeology without the restriction of working with existing systems, since few have been designed with the archaeologist in mind. Thus, the use of a GIS to provide measurements of such concepts as terrain variability, view quality, vegetation diversity, or point-to-point visibility, for example, will be discussed. The ability to compute such data, of course, may not available in most commercially produced geographic information systems, yet it is these kinds of data that are vital if geographic information systems are to be useful tools, rather than restrictive tools, for archaeological research. Archaeologists should certainly have the ability, monetary or otherwise, to influence software developers to provide necessary computer programs, and many archaeologists are rapidly gaining expertise as computer programmers themselves. Moreover, many governmental agencies employ programmers to meet the various information needs of their personnel. Hence, there is little reason why archaeologists should not have access to a GIS with capabilities tailor-made to meet their analysis needs.

The Fundamentals of Geographic Information Systems

Geographic information systems are computer-based means for assembling, analyzing, and storing varied forms of data corresponding to specific geographical areas, with the spatial locations of these areas forming the basis of the system (Tomlinson et al. 1976). The term *GIS*, as used here, is restricted to computer systems that are able to interrelate sets of data representing different geographical variables, as opposed to systems that merely manipulate or map individual files of geographical data (Rhind 1981). As Bryant and Zobrist put it, geographic information systems "seek to capitalize on the synergism inherent in being able to automatically compare a variety of socioeconomic, environmental, and land use data sets for the same point on the ground" (1977:120).

Virtually any type of geographically distributed information from any source can potentially be encoded in computer-compatible form. By using a GIS it is possible to extract information from digital geographic data bases, manipulate the data, derive new data, and analyze this information to propose solutions to problems. Thus, geographic information systems are able to transcend the role of merely processing and displaying information; they also can be incorporated into the analysis, interpretation, and problem-solving aspects of research in geographically distributed phenomena and processes (Hasenstab 1983a).

Many types of geographically distributed data can serve as the primary information portion of a GIS: elevation data; river and stream locations, vegetation patterns and soil types (which might be derived from satellite remote sensing), known archaeological site locations, and regions of planned construction or development are examples. At its simplest, a GIS can be used to retrieve spatially distributed information that is encoded in data bases for a specified coordinate point, such as the locus of a small archaeological site.

Such a procedure, however, does not fully utilize a central capability of geographic information systems—the ability to derive new information beyond that originally encoded in the data base (Collins and Moon 1981). For example, from interrelationships between known points of elevation in the data base it is possible to estimate, at any locus, values of slope, aspect, and a variety of local relief and terrain variability measures, or major drainage basins can be defined using the same data (Monmonier 1982:76–79). Points of vantage, such as hilltops and ridges, can also be determined (Kvamme 1983a). From a digital hydrology net, distances to nearest seasonal or permanent streams can be computed, and from digitized vegetation data, distances to a specified plant community (Lee et al. 1984), complex indices of vegetation diversity, or even local caloric potential can be measured. Listings of nearest neighbor sites and distances can be obtained, as well as the distance to a central place village from a data "layer" containing known archaeological site locations.

An important benefit of the data-generating capabilities of geographic information systems is that information that was previously impossible to obtain owing to the sheer number of required calculations can be derived. Maximum view distances, measures suggesting shelter or view quality, and least-effort travel distances are all potential information classes that illustrate this property. The next section discusses in greater detail the nature of these various *analytical surfaces.*

GIS Analytical Surfaces

A central GIS concept is that of *analytical surfaces,* which refers to the individual "layers" or data planes of information in a geographic data base (National Research Council 1983:41–43). Primary sources of information necessary for the construction of a GIS must be encoded in computer-compatible form. For regional archaeological research, primary information might include environmental data, such as elevation contours, river and stream locations, and vegetation and soils types, as well as cultural data, such as known archaeological site locations, archaeologically field-inspected regions, access roads, and areas of planned development or impact.

It is possible to obtain through the U.S. Geological Survey or private companies many types of geographical data, particularly regional elevation data, already in digital form and on computer tape. For example, digital terrain tapes, which were originally produced by the Army Map Service (now the Defense Mapping Agency), are available at low cost from the U.S. Geological Survey (National Cartographic Information Center 1980). The digital terrain tapes were produced by digitizing the elevation contours on 1:250,000 scale topographic series maps, and they are available for the entire United States (Doyle 1978:1484). As might be expected, these data are somewhat crude owing to the scale of the original map sources, and recent studies (Stow and Estes 1981) point to inaccuracies in the resulting elevarion surfaces (e.g., small ridges, drainages, and canyons are underrepresented).

As an alternative, the USGS is currently producing highly accurate digital elevation models (DEMs) that are obtained through digitization of 1:24,000 scale topographic maps (Doyle 1978:1484). Not only the elevation data but also other classes of planimetric information, including hydrologic and cultural data, such as road networks, are available for these maps. A limitation of this data source is that only a small percentage of the quadrangles across the country have been digitized to date, although the USGS ultimately plans to digitize all the 1:24,000 scale maps. For a particular study region, high-quality elevation and hydrologic data, two of the most important sources of information for archaeological locational studies, may already be available in digital form. It is unlikely, however, that other sources of information, such as vegetation and soil data, will be available in digital form, and archaeological data certainly will not be available. As a result, it is often necessary to digitize these data electronically.

A common digitizing procedure utilizes a digitizing tablet and cursor (Monmonier 1982:7; Rogers and Dawson 1979). With these devices, such pictorial information as elevation contour lines or stream courses are manually traced and encoded in computer-compatible form (Figure 10.1). The tablet may contain as many as a million x,y coordinates per square inch (Calcomp 1983); as the lines are being traced they are electronically converted to corresponding x,y coordinates that the computer is able to utilize. This procedure is, of course, somewhat labor

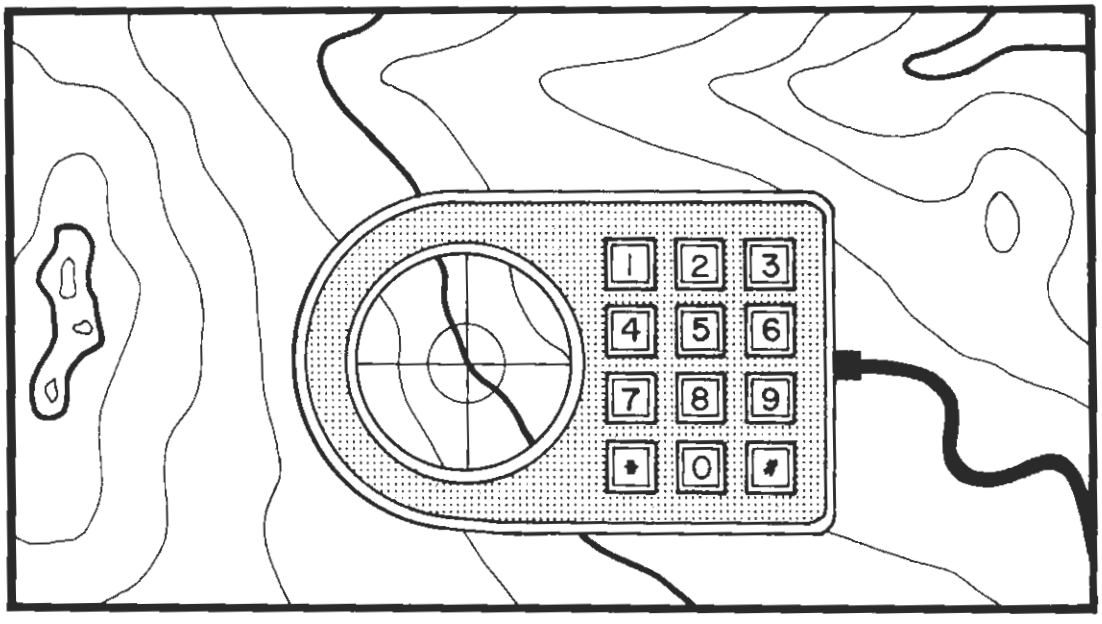


Figure 10.1. Manual digitizing of contour lines through use of a cursor and digitizing tablet. Pictorial map information, affixed to the tablet, is converted to x,y coordinates by manually positioning the cross hairs of the cursor over the intended point and pressing a button. The keys on the cursor control different functions or allow entry of category codes.

intensive. For example, digitizing the elevation contours on a typical USGS 7.5-minute quadrangle can take anywhere from one to six (or more) person days, depending on the complexity of the terrain. State-of-the-art digitizing technology utilizes optical scanners to digitize complex pictorial information in seconds (Leberl and Olson 1982), but this equipment can be very expensive.

The primary data are usually derived from traditional maps, but other sources, such as preclassified or interpreted remotely sensed digital satellite images, can be used (Shelton and Estes 1981; see below). However they are acquired, the several primary surfaces of digital information that the GIS needs are encoded and stored in the initial data base (Figure 10.2). Computer programs then are able to utilize these primary data to derive secondary information that often is more useful than the primary data (Collins and Moon 1981). For example, slope estimates, aspect estimates, or distances to nearest drainages might be derived (from elevation and hydrology surfaces, respectively) and stored as new and distinct analytical surfaces (Figure 10.2).

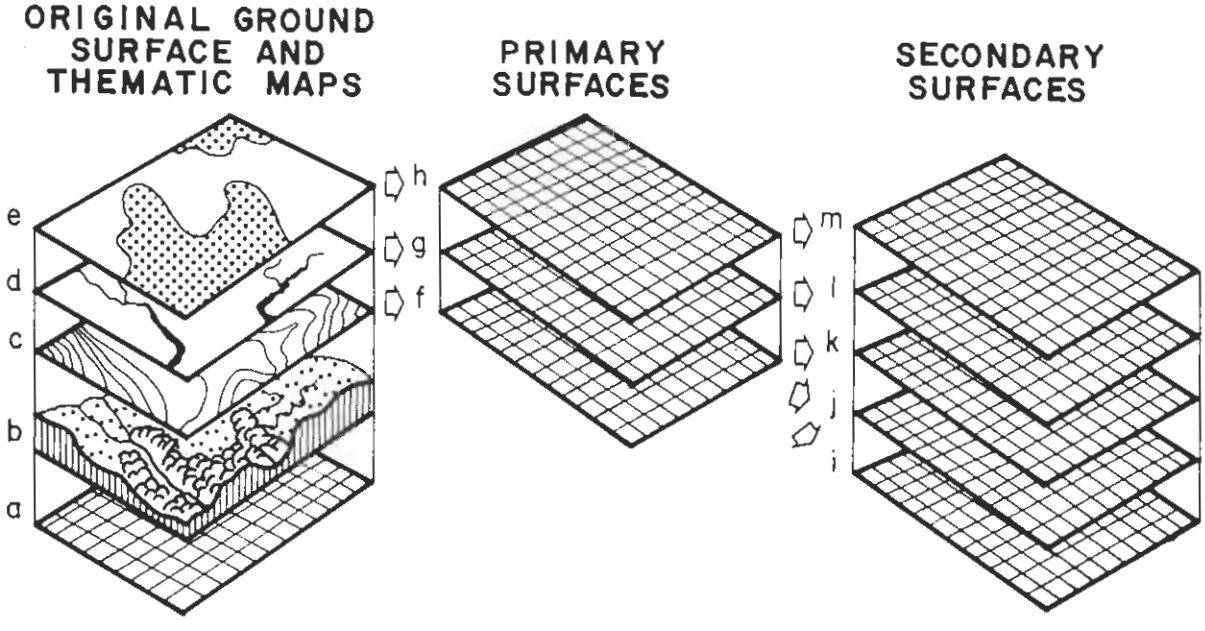


Figure 10.2. Construction of a GIS. From the original land surface (b), various thematic maps are produced, such as elevation contours (c), hydrology (d), and forested areas (e). These maps are digitized and converted to primary layers in a GIS representing an elevation surface (f), a hydrology surface (g), and a forest location surface (h), which are all referenced to a reference grid, such as the UTM grid (a). From the elevation surface such secondary surfaces as slope (i), aspect (j), and local relief (k) might be obtained. The hydrology surface might provide a secondary surface showing distance to nearest drainage (l), and the forest location surface might yield a surface showing distance to nearest forest (m).

A basic principle of geographic information systems is that the users provide the system (through digitization or other means) with the minimal information that it needs (primary layers). The GIS itself subsequently derives secondary sources of data by means of various software techniques. Both primary and secondary surfaces can then be used for analytical or display purposes. The specific ways in which these data are utilized, however, depend on the nature of the particular GIS.

GIS Types

There are two fundamental GIS designs. A vector-based GIS, such as the Department of the Interior's MOSS (Lee et al. 1984), stores data as a series of points, lines, or polygons that are used to identify discrete features that typically occur on traditional maps (*vector* is another word for a line between two points). A cell-based (sometimes called raster-based) GIS, such as the Department of the Interior's MAPS (a subsystem of MOSS), superimposes a regular grid containing rows and columns of cells over the region and assigns a numeric value to each cell (Figure 10.3). Each design has certain advantages and disadvantages in terms of archaeological locational analysis and modeling.

Vector-Based Geographic Information Systems

Vector-based geographic information systems accommodate information digitized as points, lines, or polygons (i.e., mappable data; Figure 10.3). Computer storage requirements for this information are minimal since only the coordinates of digitized points (points along line or polygon boundaries) are stored. A vector-based system is suitable for cultural resource *information management* since various mappable entities—archaeological sites, site boundaries, surveyed regions, and archaeologically sensitive zones—are easily retrieved and displayed, as are other types of discrete map information (e.g., specific soil type locations). A vector-based GIS can also be used for the display of very simple site location models that are based on a one-to-one correspondence between the locations of sites and discrete categories of information, such as plant community or soil type locations (Thompson 1978; see also Cordell and Green 1983). For example, if a site location model suggests high site density in piñon-juniper settings, a vector-based system can easily present a series of polygons showing the locations of high-site-density piñon-juniper zones.

Although vector-based geographic information systems can be used to manage and display discrete classes of map data, these systems are unsuitable for many of the analysis and modeling techniques described in earlier chapters. In analysis and modeling contexts, systematic measurements or observations of environmental or other features are required (e.g., every 50 or 100 m across a region of study). In other words, spatially contiguous values of the data are necessary. In vector systems such information is not available; data values are present only at point, line, or polygon boundaries, which constitute only a very small portion of any region. This short-

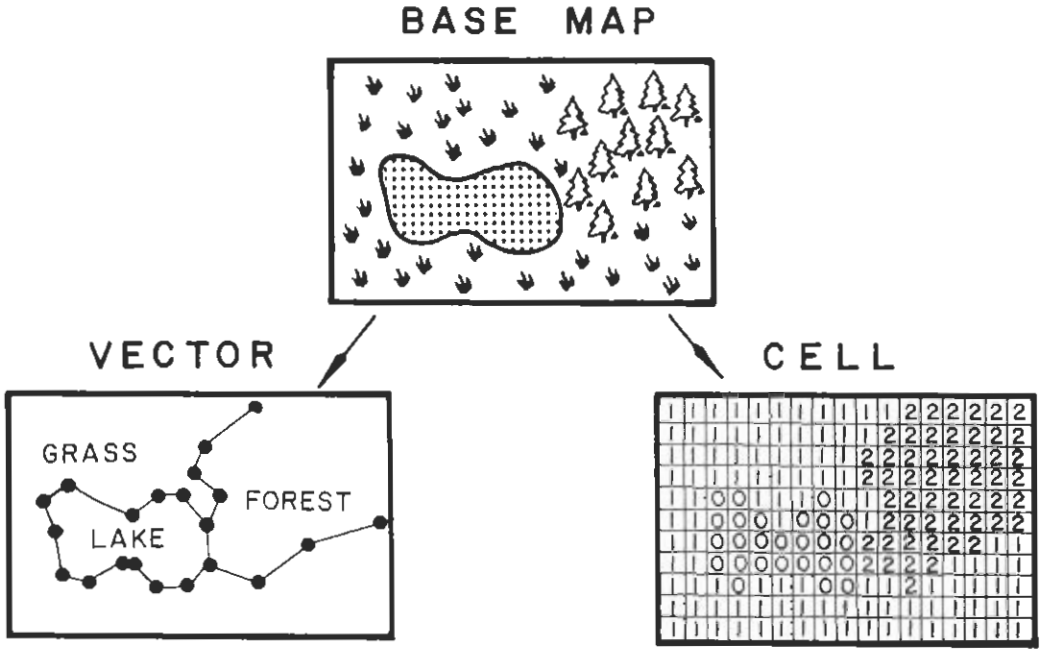


Figure 10.3. Characteristics of map information in vector-based and cell-based geographic information systems (after Lee et al. 1984).

coming is further illustrated by the fact that even if continuously varying map information is available in digital form, such as elevation or slope values, such data must be transformed to line data by contouring or by categorizing the continuous measurements into discrete classes (e.g., level vs steep slopes) to be handled by a vector GIS.

Cell-Based Geographic Information Systems

With a cell-based or raster GIS, both categorical and continuous map information can be incorporated. Since a grid is superimposed over the entire region, each analytical surface is composed of rows and columns of grid cells, each cell corresponding to a fixed area in real space and each containing a value for that area (Figure 10.3). For example, an elevation surface would contain an elevation in each cell representing the height of the ground; a slope surface would contain a slope measurement in each cell; and a nominal-level surface, such as a representation of plant community locations, would contain a unique value in each cell, with each value corresponding to a specific plant community class. Since a value must be stored for each cell for each analytical surface, cell-based geographic information systems typically require large amounts of computer storage. Owing to the gridding or *rasterization* of features, the quality of display of information can suffer to some

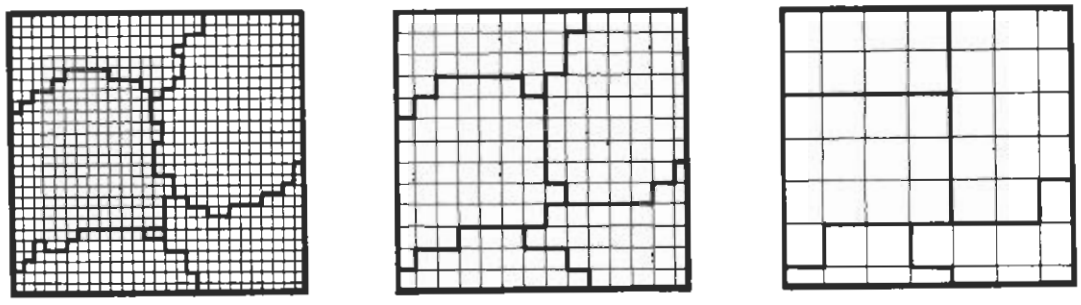
extent (Figure 10.3), although this depends on the resolution (size) of the cells and of the display device (see below). These difficulties are decreasing, however, because mass storage and high-resolution display devices are rapidly becoming available at low cost.

Since cell-based geographic information systems can accommodate continuously varying and categorical information, treating each as a surface of contiguous values, and since they can easily derive and store many types of new data of relevance to archaeological inquiry over entire study regions, this type of GIS is well suited for archaeological locational analysis and modeling research. Additionally, with a cell-based GIS each analytical surface, regardless of type, can be treated as an "image" (referred to as a *pseudo-image* in image analysis). This means that the researcher can make use of the large number of available image analysis, manipulation, and classification techniques (see Chapter 9 for an overview of some of these), as well as a host of image-processing software packages (see Kohler's brief overview later in this chapter). The following sections focus on cell-based geographic information systems since they are better suited for the archaeological analysis and modeling approaches discussed in this volume.

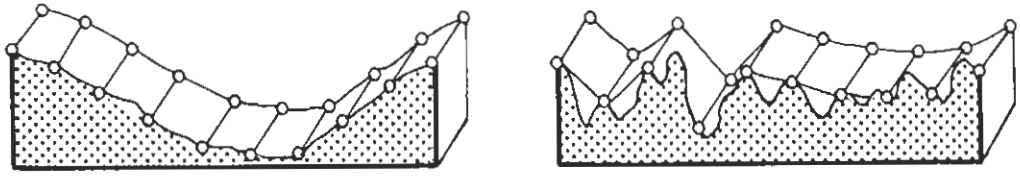
GIS Issues

Several issues in GIS research are of importance to archaeological modeling applications. One issue is that of cell size in a cell-based GIS (Wehde 1982). The size or resolution of the cells is extremely important because it determines the nature and quality (accuracy) of the features that can be analyzed. For nominal-level features, such as vegetation community locations, a large grid may severely misrepresent the true shapes and sizes of the categories, which may result in inaccurate border and area estimates (Figure 10.4a). For continuous data, such as an elevation surface, large cells tend to smooth features of the terrain; small ridges, canyons, or drainages might be underrepresented, less pronounced, or even invisible on the gridded surface (Figure 10.4b). An additional result is that any surface derived from such an elevation layer (e.g., slope, aspect, relief, and ridge identification; see below) will also be smoothed.

Although small cell sizes may portray various features more accurately, an important consideration is that computer storage requirements increase geometrically with decreased cell size. For example, to store the information from a typical 7.5-minute USGS map gridded in cells 100 m on a side (about one-sixth of an inch on the map) would require about 15,000 cells per layer of data; cells 50 m on a side (about one-twelfth of an inch on the map) would require about 60,000 cells per layer. Thus, some balance must be struck between cell resolution and storage requirements. It should be emphasized, however, that small cell size does not necessarily guarantee accuracy. It is technically possible, for instance, to increase the resolution in any data plane (say from 200 m to 30 m on a side), but if the data were initially



A



B

Figure 10.4. Effects of cell size in cell-based geographic information systems. (A) Sizes and shapes of discrete classes can become severely distorted. (B) For continuous data, such as an elevation surface, the cell size may be adequate to display terrain features (right) or it may be inadequate, resulting in a smoothed surface (left).

encoded at the grosser level of resolution the final result would offer no increase in accuracy.

Another important concern in a multilayered GIS pertains to registration of the individual layers (National Research Council 1983:42). One must be absolutely certain that a coordinate point in one layer lines up in real space with the same coordinate point in other layers. For a cell-based GIS this means that the borders of each cell in each layer must coincide (within acceptable limits) with the borders of each cell in the other layers. This is a particular problem when combining data from such diverse sources as aerial photographs, remotely sensed images, and a variety of map projections and scales. A wide variety of procedures for registration of multiple data sources can be found in a number of standard image-processing sources (in particular, Moik 1980:187–198; Schowengerdt 1983:99–116).

The computational algorithms thar a particular GIS employs should not be taken for granted. Many commercially available geographic information systems make use of procedures that are crude and unsuitable for archaeological locational research dealing with relatively small study areas or site-specific microenvironmental characteristics. This stems, in part, from the fact that many geographic information systems are designed primarily for low-resolution uses, such as county or statewide administrative planning and mapping needs. Technical manuals that accompany most GIS packages usually provide some information about the computational approaches that the system uses.

GIS Algorithms

Primary Surfaces

One of the most important primary surfaces in a geographic data base, at least for archaeological analysis and modeling purposes, is the *elevation surface* because this surface represents the form of the terrain in a region. A wide variety of terrain features, such as slope, aspect, local relief, terrain variability measures, hilltop or ridge "vantage" locations, view quality, and shelter quality measures, can potentially be derived from this surface. This surface is also one of the most difficult to construct unless it can be obtained preestablished from some outside sources (as noted earlier). In order to portray more clearly some of the software mechanics behind a geographic data base, one way (out of many possible ways) to consrruct such a surface will be described.

An elevation contour can be represented as a series of lines between digitized points, as in Figure 10.5a. In constructing a cell-based elevation surface where an elevation value is available for every locus (cell) in the GIS region, the first step might be to place these digitized points in appropriate cells (Figure 10.5b) and connect the cells between the points (Figure 10.5c) to yield a gridded or *rasterized* image of an input contour map (Figure 10.5d). In this layer (in computer terms, a two-dimensional array) cells (array elements) that contain a contour possess the elevation value of rhat contour, while other cells contain a zero. This array of elevation contours must be transformed to an *elevation surface* in which every cell contains an elevation value.

Interpolation routines are used to provide an initial estimate of the elevation at every poinr (cell) where rhe elevation is unknown (the zero cells). There are literally hundreds of interpolation algorithms (e.g., Delfiner and Delhomme 1975; Rhind 1975; Yoeli 1975:360–366); some provide more accurate estimates of unknown elevations but use larger amounts of computer time, others provide less satisfactory estimates but require less computation. The amount of time used by a particular algorithm can be an important consideration given the large number of estimates typically required for even moderate-sized regions. Two common interpolation algorithms are illustrated in Figure 10.6. The first, a column (or row) scan, which

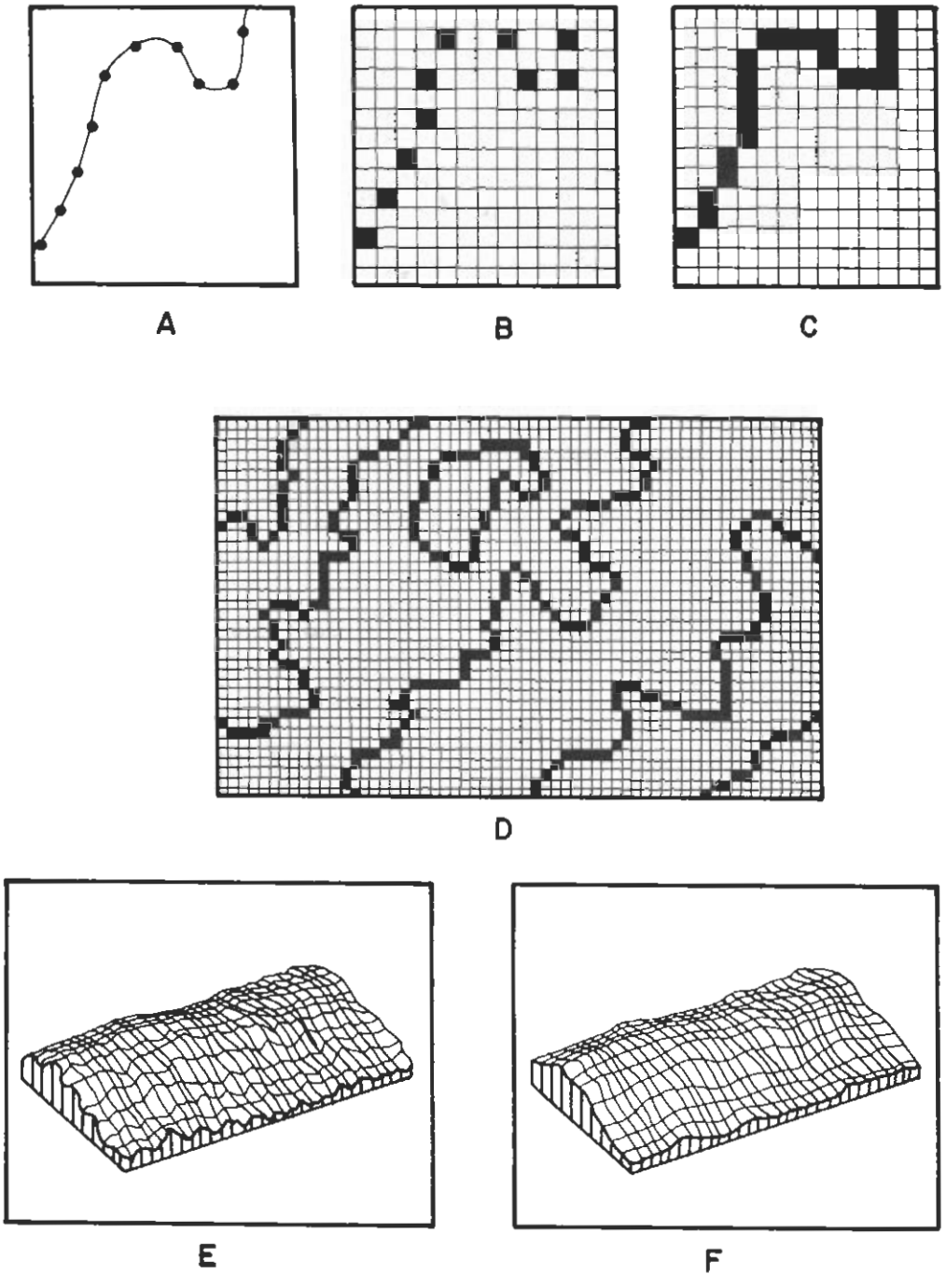
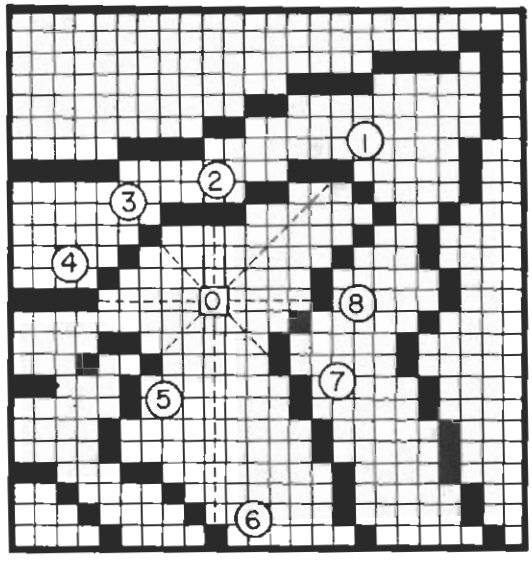


Figure 10.5. Steps that might be followed in the construction of an elevation surface. (A) Digitized points are indicated on a contour line. (B) The digitized points are placed in appropriate grid cells. (C) The cells between the digitized cells are filled in. (D) In a gridded or rasterized version of the original contour map, contour line cells contain the elevation value of the contour; empty cells contain a zero. (E) The initial surface of interpolated elevations is "noisy." (F) The final elevation surface is smoothed.



$$(A) \quad ELEV(0) = \frac{ELEV(6) - ELEV(2)}{DIST(6.2)} \times DIST(2.0) + ELEV(2)$$

$$(B) \quad ELEV(0) = \frac{ELEV(5) - ELEV(1)}{DIST(5.1)} \times DIST(1.0) + ELEV(1)$$

Figure 10.6. Elevation interpolation algorithms. (A) A column (or row) scan technique searches along columns (or rows) only and linearly interprets unknown elevations (0) between points of known elevation (2, 6). (B) In the second technique an algorithm searches in eight directions (1–8) from a locus (0) of unknown elevation, finds the line of steepest change (1, 5), and linearly interpolates an elevation at the locus. Note that this procedure follows manual interpolation techniques more closely than the scan procedure.

searches for known elevations only along a given column (or row), requires little effort to compute but may not offer a good estimate of the unknown elevation in some situations. The other, which searches in eight directions for known elevations and uses the line of steepest increase as a basis for interpolation, usually produces more accurate results (in fact, closely mimicking manual interpolation techniques) but takes greater computational effort and therefore more computer time. (This comparison illustrates the point made above about the importance of examining computational procedures.)

The outcome of the interpolation routine is an initial elevation surface (Figure 10.5e), which can be extremely "noisy," containing many small, artificial peaks and valleys. Because each elevation in this surface is interpolated independently, each

elevation might be estimated from known values that are quite different from the values used to estimate adjacent elevations, resulting in some disparity between adjacent elevation estimates. The final step in creating an elevation surface, called *smoothing*, attempts to remove this noise by providing a better elevation estimate at each location (Allan 1978:1518; Monmonier 1982:65–66). This smoothing process (which is distinct from the detrimental smoothing caused by large cell sizes) recognizes that elevation estimates in adjacent cells, because of their proximity or spatial autocorrelation, should also be good estimates of each cell's elevation. A final estimate in each cell is therefore typically accomplished by taking a weighted average of each cell's elevation and the elevations in the adjacent cells (with most weight being given to the current cell). The more familiar smoothing in one dimension is illustrated in Figure 10.7a, while two-dimensional smoothing is shown in Figure 10.7b. The resulting surface, without the artificial peaks and valleys, is illustrated in Figure 10.5f.

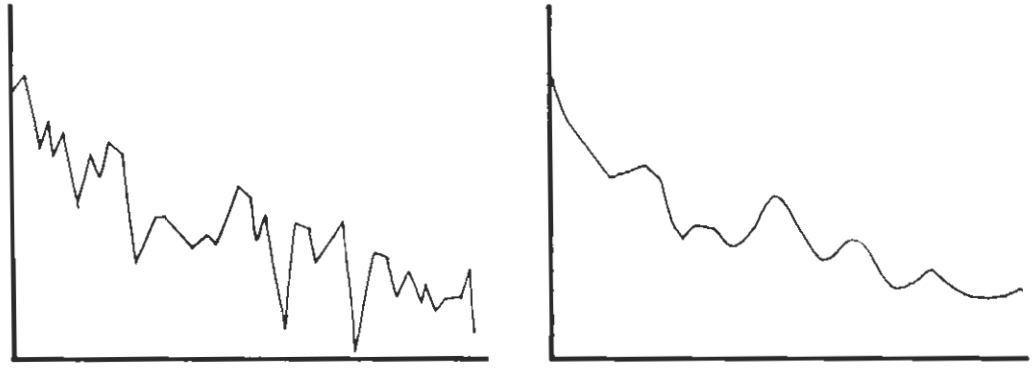
Orher primary surfaces are somewhat easier to obtain (if not already available commercially). For a hydrology net, the stream locations are digitized in much the same way as elevation contours (Figure 10.8a). The digitized streams are then placed in grid cells to form a rasterized image of the hydrology net (much like the rasterized elevation contours in Figure 10.5d). The streams, however, might be coded to reflect permanent or seasonal water (Figure 10.8b) or Strahler order ranks (Figure 10.8c; see Chapter 8 for a description of the Strahler order ranking system).

Rasterization of polygonal areas, lines, and points, which are used to describe discrete classes of information, such as vegetation communities, soil types, archaeological site locations, and archaeologically field-inspected regions, is fairly straightforward. Digitized polygons are merely transformed to a gridded version of the polygons (Figures 10.3 and 10.4a) using various polygon-fill routines (MacDougall 1971:117–126; Monmonier 1982:68–73). Polygon cells that represent a particular class are assigned a unique identification number.

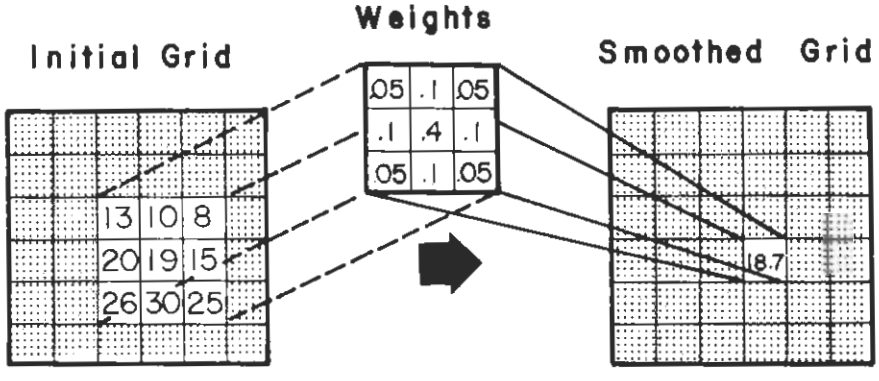
Secondary Surfaces

An infinite number of secondary analytical surfaces of potential importance to regional archaeological research can be derived from the primary surfaces in a GIS framework. Two common types are slope and aspect. Based on interrelationships between the elevation of a grid cell and those of its nearest neighbors in the elevation surface, some algorithms (e.g., Woodcock et al. 1980) fit a least-squares plane to these elevations and find the maximum slope and the direction of maximum slope (aspect) on this plane (Figure 10.9a). Other algorithms might find a maximum, minimum, or average slope (e.g., MOSS; Lee et al. 1984).

A variety of terrain variability measures are easy to obtain from the elevation surface (see Chapter 8 for more detailed discussion of these variables). For example, local relief (maximum minus minimum elevation) can be obtained within any defined radius of a given cell (Figure 10.9b). Another terrain roughness measure is



A



$$.05(13) + .1(10) + .05(8) + .1(20) + .4(19) + .1(15) + .05(26) + .1(30) + .05(25) = 18.7$$

B

Figure 10.7. Illustration of smoothing (after Monmonier 1983). (A) Smoothing in one dimension: the original noisy trend (left) is compared to the same trend after smoothing (right). (B) Smoothing in two dimensions: a smoothed surface (right) is obtained by calculating a weighted average of the initial grid (left).

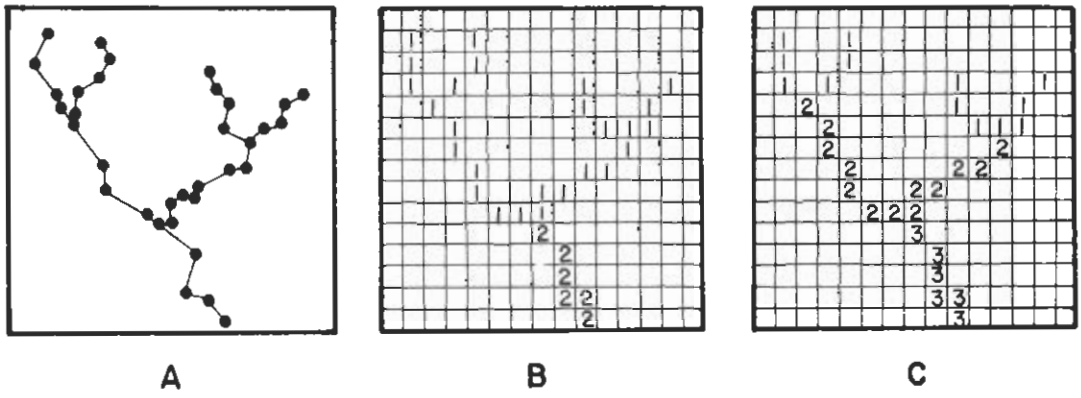


Figure 10.8. Encoding of hydrologic data. (A) Digitized points are indicated on a hydrology net. (B) The stream locations are placed in grid cells. Seasonal water might be coded as "1" and permanent water as "2." (C) Streams might also be coded according to Strahler order ranks.

termed a texture measure in image processing (Moik 1980:232). This measure finds the variance of elevations within a defined radius or "window" of a given location (Figure 10.9b): high values suggest variable or rough terrain while low values suggest level or smooth terrain. Fragmentation indices (Monmonier 1974) provide other analytical alternatives.

Hilltop, mesa edge, and ridge crest vantage locations might be defined using a variety of techniques (e.g., Kvamme 1983b). For example, the previously derived slope data plane might be used to define all level locations (e.g., those with grades less than or equal to 15 percent) adjacent to or within a certain distance of steep locations (those with grades greater than 15 percent). The elevation surface is then used to delimit those locations (cells) above the adjacent steep locations.

An angle of surrounding view, one possible measure reflecting quality of view, can be obtained from the elevation surface simply by calculating for each cell the angle that encompasses all elevations in the surrounding eight cells that are less than the current cell's elevation (Figure 10.9c). A "view catchment," another possible measure of view quality, might be calculated by fixing a 1 mi radius around each cell and calculating the percentage of cells within that radius that are visible from the current cell (Figure 10.9d; Lee et al. 1984).

More traditional catchments might be calculated using a nominal-level vegetation layer. Given a fixed catchment radius around each cell (Figure 10.9d), the proportion of various plant communities within that radius can be obtained and stored in separate derived layers. Alternatively, some index of vegetation diversity or complexity or some estimate of caloric potential might be calculated.

Search and distance-measuring routines can be used to derive a variety of analytical surfaces; the MOSS-MAPS system, for example, has several (Lee et al. 1984). The nearest specified water type (e.g., seasonal, permanent, or a stream of

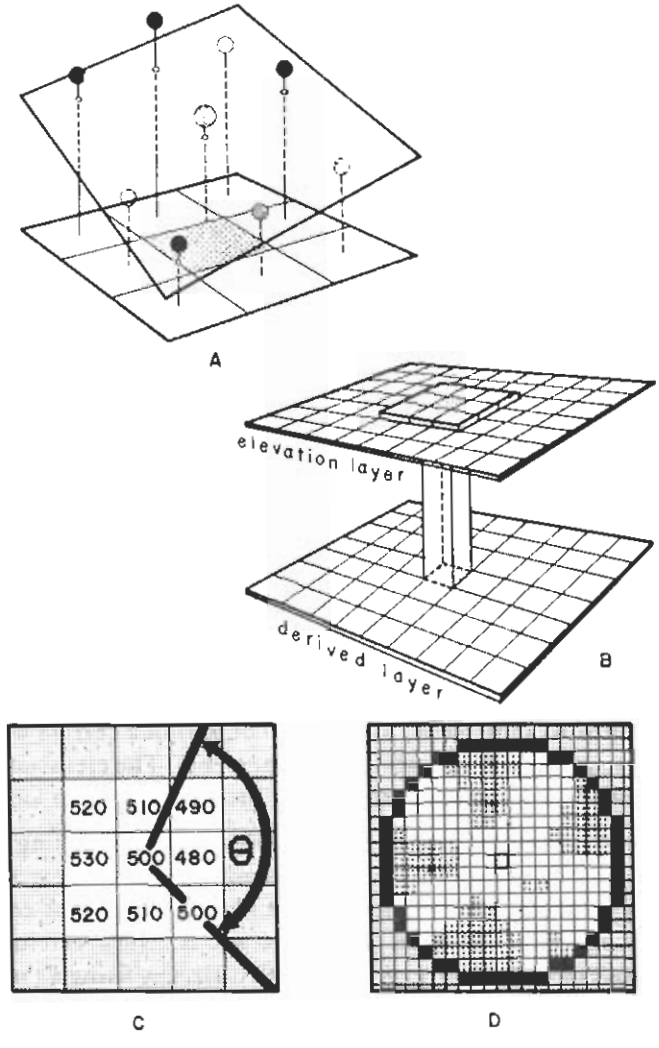
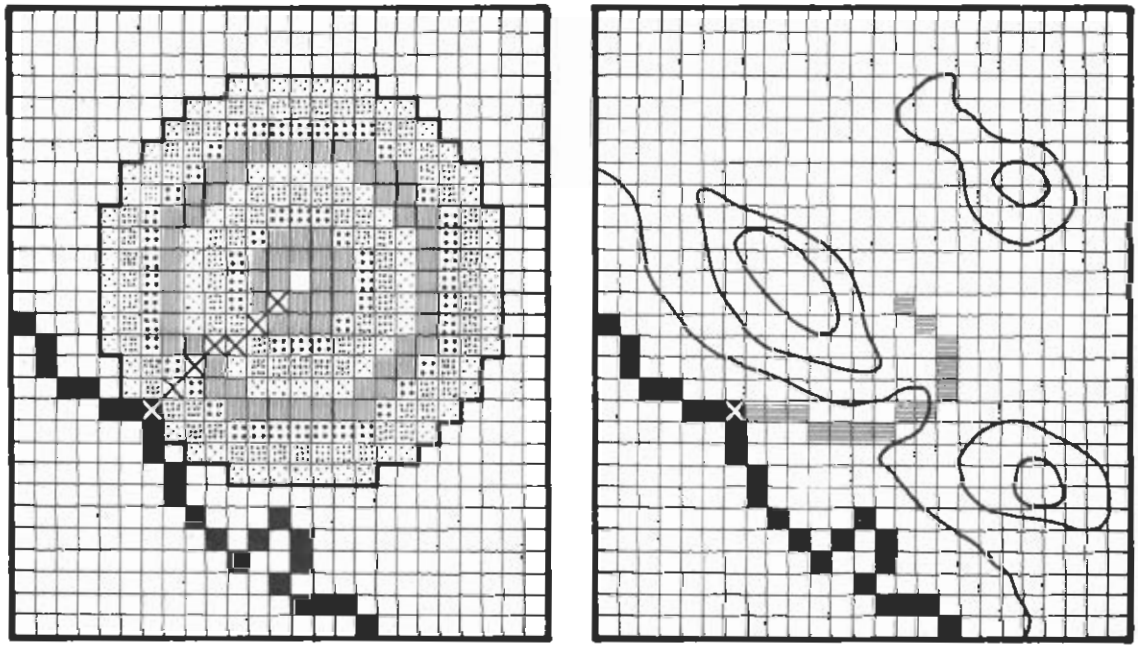


Figure 10.9. Examples of various computational algorithms. (A) A least-squares plane might be fitted to an elevation (the central sphere and shaded cell) and its eight nearest neighbor elevations. The maximum slope on this plane might be calculated, along with the direction of maximum slope (aspect). (B) Local relief might be calculated as the range in elevations in a three-by-three window around a current elevation. Alternatively, the variance of the elevations might be calculated to derive a texture measure. (C) An angle of view could be calculated in an elevation surface. (D) A catchment radius can be fitted around a cell. Areas or percentages of the feature of interest can be calculated within the radius.

specified Strahler order rank) might be located from a primary hydrologic net, for instance, and the horizontal Euclidean distance could be calculated for each cell (Figure 10.10a). In conjunction with the elevation surface, the vertical distance to the same drainage type might also be obtained. If hilltop, mesa edge, or ridge crest vantage points are already defined, search procedures can be used to obtain a distance to nearest vantage within each cell or, using a vegetation community surface, the distance to a specified plant community. Linear distances, however, might not be the best measure to use in site location studies (Ericson and Goldstein 1981); because there often are obstacles to cross, people do not normally follow straight paths. If appropriate software is available, and definitions of "effort" can be made (see Turner 1978), least-effort travel distances might be estimated instead (Figure 10.10b).

Geographic information systems can accomplish many of the same tasks using "cultural" variables as they do for environmental ones. For example, if central place sites are defined in the data base, then distance from each cell to the nearest central place can easily be generated. Similarly, based on the locations of known archaeological sites, various orders of nearest neighbor site distances can be calculated.

These examples illustrate the kinds of phenomena one might potentially investigate in site location studies through the use of GIS techniques. Such investigations are limited only by our ability to innovate and be creative (and by CPU and storage requirements)!



A B

Figure 10.10. Illustration of distance calculation techniques. (A) To obtain linear distance the computer scans from a current cell with search radii of increasing length until the feature of interest is encountered. (B) Measurement of least-effort travel distance might consider paths that avoid hills.

Geographic Information Systems and Remote Sensing

As discussed in Chapter 9, the potential of remotely sensed data for a number of applications in archaeology and elsewhere is beyond question. Recently, a number of remote sensing specialists have noted that geographic information systems "have significant potential to facilitate use of remotely sensed data" (Shelton and Estes 1981:395). A key problem in remote sensing, for example, is that remote sensor imagery is usually geometrically distorted; for these data to be useful in applied contexts the interpreted information must be transferred to a standard geometrical base or georeference (Steiner and Salerno 1975:622). Tilted or oblique satellite images must be rectified to a horizontal reference plane. The rectified image then must be geometrically corrected into a particular map projection, such as longitude and latitude or the UTM system. The importance of these tasks is recognized by the Jet Propulsion Laboratory (JPL), a center of state-of-the-art remote sensing and image processing. McLeod and Jafek (1984:75–76) note that

> Perhaps the most prodigious technology introduced by the lab is that of the geographic information system, which co-registers and analyzes a virtually limitless supply of sensor data types, and then relates them to key geographical questions within a given region. At one pole of true state-of-the-art image processing, GIS is the reverse of the imaging technique that solely enhances immediate visual recognition within a particular scene or image data set. Rather, GIS is JPL's answer to the need for analysis of unmanageably large data bases and the need to make responsible decisions about them.
>
> . . . Each image is first entered into the data base and geometrically corrected before being registered to the "planimetric base" or system of data planes. Each image plane is again referenced to one or more georeference planes. The user is thus able to manipulate data from several sources which, despite their original disparity, are referenced to a common base.

Since geographic information systems interrelate multiple geographic data sets that are tied to specific locations, it is clear that the JPL system, although it primarily uses remotely sensed data, meets this definition.

There are other reasons why geographic information systems and remote sensing should logically be linked. In recent years various forms of ancillary data, such as digital terrain models (see above), have been incorporated into remote sensing applications. During a project that developed classification models for forest cover type based on remotely sensed spectral data, for example, it was discovered that incorporation of ancillary terrain data, such as elevation, slope, and aspect, significantly improved the classification accuracy of the predictive models (Hoffer et al. 1975; Strahler et al. 1978; Woodcock et al. 1980). Although spectral signatures could distinguish plant cover types to a fair extent by themselves, it was found that the distributions of many plant groups were also related to such factors as ground steepness, aspect, and elevation (Hoffer et al. 1975; Strahler et al. 1978:930), variables that were not readily obtainable from the remotely sensed imagery. By merging digital terrain models and the remotely sensed spectral data into a single analytical data set, not only could the elevation data be obtained, but through

various software techniques, estimates of slope and aspect could be derived, allowing more powerful predictive models to be developed. The success of these approaches has led to applications using more varied forms of ancillary data combined with remotely sensed imagery. Missallati et al. (1979) combined detailed geologic map data, aeromagnetic data, and radiometric data (all digitally encoded) with Landsat spectral information to develop predictive models for uranium exploration. Loveland and Johnson (1983) combined remotely sensed data with digital terrain data and digital soil survey, land ownership, and pumping plant location data to develop predictive models to evaluate irrigation agriculture. This project showed, as Loveland and Johnson put it, "the flexibility of remotely sensed and other spatial data as input for predictive models" (1983:1183).

Geographic information systems are potentially useful for manipulation of geographic data regardless of their source. Recently, this fact has generated considerable interest in remote sensing circles (see Shelton and Estes 1981 for an overview). A new perspective has arisen that suggests that the focus of research should be on the *region* under investigation (rather than on particular sources of data) and that *all* relevant sources of information, regardless of type or derivation, should be sought for input into the regional GIS. Potential data sources include traditional thematic maps and a variety of remote sensor inputs. In this context, the GIS treats each analytical surface, regardless of source, as simply another data plane. The GIS is able to facilitate manipulation, analysis, and modeling of these varied data types, treating information sources individually or in combination.

The importance of incorporating remotely sensed data into comprehensive geographic information systems is summarized by Shelton and Estes (1981:417):

> the full potential of remote sensing cannot and will not be achieved without continued and expanded efforts to adapt the technology to the evolving needs of users around the world. To the extent that geographic information system designs reflect those needs, GIS design ought to be a relevant concern in the development of new satellite systems and in establishment of institutional arrangements for processing, formatting, and disseminating the products of remote sensing.

As a final caveat, however, they note that geographic information systems represent an evolving technology. Since remote sensing can contribute to the development of a GIS, e.g., by providing varied forms of data input, they conclude that full acceptance of *both* of these technologies "is dependent on realization that the potential of each technology will not be achieved until they are integrated."

The Potential of Geographic Information Systems for Regional Archaeological Research

GIS techniques may potentially contribute in a number of ways to regional archaeological site location research and modeling, and these techniques may have

numerous applications to cultural resource data base management as well. Some of these potential applications were suggested in a foregoing section on fundamental concepts; the following sections will elaborate on these suggestions and add several additional ones.

Spatial Data Management

A GIS can consolidate and merge many and diverse forms of geographically distributed information into a single data base. This is perhaps the most obvious application of GIS technology to regional archaeological research. Since archaeological data inherently are geographically distributed, they are well suited to a GIS context. Varied forms of archaeological data, such as archaeological site locations, site types, regions that have been field inspected, and cultural resource sensitive locations, can be merged into a single data base, together with varied sources of environmental and other geographically distributed data. Sources of information can be as diverse as traditional topographic maps, thematic maps (soils, vegetation, geology), aerial photographs, and remotely sensed spectral data (Kvamme 1986; Parker 1986).

In a regional geographic data base established for the explicit purpose of developing, testing, and applying predictive archaeological locational models in southern Arkansas, Scholtz (1981; see also Parker 1985) utilized a cell-based format containing 3479 cells, each representing an area of 4 ha (200 m sq). Fifteen biophysical variables, including soil type, elevation, slope, and distances to streams of various orders, were measured in each cell. Once the data were measured and formatted within a single computer data base, an exceedingly powerful tool was established for investigating environmental patterning exhibited by the locations of known sites and for formulating and mapping the results of archaeological prehistoric and historical locational models.

Hasenstab (1983b) developed a GIS for archaeological predictive modeling in the Passaic River Basin of New Jersey. This data base was established by electronically digitizing a wide variety of conventional maps and aerial photographs. Environmental data included soil type, landform, slope, drainage, agricultural potential, current land use, degree of disturbance, type of modern development, and distances to the nearest major river course, to confluences of major rivers, to tributaries, to confluences of tributaries with major rivers, and to major wetland zones. Management data included the location of known prehistoric and historical archaeological sites, a gross river basin division, USGS quadrangle reference, and locational coordinate information. Most of these data were generated from other digitized sources; the information was stored in 4306 georeferenced cells, each representing an area of approximately 1.15 acres.

Digital terrain tapes were used as the basic data source in a western Colorado study that attempted to model prehistoric archaeological site locations (Kvamme 1983b). Six secondary surfaces, representing slope, aspect, angle of view, local relief, vantage locations, and distances to nearest point of vantage, were generated from

the initial elevation surface for each of 5000 cells, which measured 100 m on a side. Stream courses were manually digitized, and the stream locations, together with horizontal and vertical distances to nearest streams, were included in the data base, as were the locations of known archaeological sites. Other secondary surfaces, in the form of various probability surfaces of archaeological site presence (based on various combinations of variables), were also generated from these data.

The Granite Reef Archaeological Project (Brown and Stone 1982) made extensive use of a GIS for management of the project's data and for purposes of spatial analysis and archaeological modeling. The Granite Reef project encompassed a huge area of west-central Arizona, more than 12,000 mi^2. A variety of basic environmental data was encoded for cells measuring 1.16 mi on a side, including elevation, slope, basin divides, aspect, major watersheds, geologic classes, soil classes, vegetation classes, seasonal precipitation, and elevation-adjusted temperature extremes. Encoded archaeological data included the locations of regions surveyed by archaeological field teams and a variety of site types, ranging from habitation sites to lithic scatters, sherd scatters, rock rings, rockshelters, rock art, and prehistoric trails. Based on various arguments and notions about the relative importance of each of the environmental factors to the prehistoric occupation of the region, the GIS was used to develop a number of prehistoric land-use models that were weighted composites of the basic environmental data.

Regional GIS data bases for a southern Federal Republic of Germany study area and a southern Colorado study area are described by Kvamme (1986; also see Chapters 7 and 8). These geographic information systems have similar characteristics in the nature of rhe data planes that were established and in their purposes: archaeological locational modeling. Both systems include such data as elevation, slope, aspect, and measures of local relief, view quality, vantage locations and distances to nearest vantage, and shelter quality, along with the complete hydrology network, horizontal and vertical distances to streams of various Strahler order ranks, and the locations and types of archaeological sites (approximately 200 sites in the German data base and 1200 sites in the Colorado data base). The German GIS contained nearly 80,000 cells, each encompassing 1 ha, and the Colorado GIS contained approximately 230,000 quarter-hectare cells. Both systems were used to establish archaeological locational models based on logistic regression probability functions; these models were stored as separate GIS surfaces.

In the above geographic information systems various sources and combinations of management and environmental data, such as archaeological information about a particular site and its environmental properties or scaled maps of any surface or combination of surfaces, can be retrieved. One of the chief uses of the geographic data bases in all of the above studies is to examine and test environmental hypotheses about archaeological site locations and to develop various settlement pattern models, including those used for the explicit purpose of prediction.

Generation of New Data

The use of a GIS makes it possible to derive new data and to explore new variables and measurement concepts. The ability to derive new data from primary information initially encoded in a geographic data base was discussed at length in an earlier section. The speed and accuracy of computers not only allow vast quantities of information to be generated but also permit extremely complicated and time-consuming measurements to be performed. The large numbers of measurements of elevation, slope, aspect, distance to water, etc., that can be produced demonstrate in part the tremendous workload capabilities of computers. Another example involving the computation of point-to-point visibility through the use of an elevation surface illustrates the complexity of calculations that can be performed. From a given location (grid cell) of known elevation, one algorithm first approximates the straight-line path through the reference grid of cells to the desired point or grid cell, which is also of known elevation. If cells in the straight-line path contain an elevation higher than the highest of the two end-point cells, a determination of *no visibility* is immediately made; if the intervening cell elevations are all lower than the lowest of the two end-point cells, a determination of *visibility* is immediately made; otherwise the standard point-slope formula is invoked to determine the equation of the line-of-sight between the elevations of the end-point cells. In this third case, the actual elevation for each intervening cell is compared with the computed line-of-sight elevation at that cell locus to determine if visibility is blocked (Creamer 1985). Performing this procedure by hand between only two locations would be incredibly time consuming. Performing such a procedure between many hundreds of hilltops is impossible without the use of a computer.

Computer Cartography

Within a GIS it is easy to display information using computer graphic/cartographic techniques. Advances in computer graphics and cartography (e.g., Edwards and Batson 1980) allow maps to be produced rapidly and accurately, incorporating uses of color, shading, and three-dimensional perspective that are unavailable in traditional cartography. The flexibility of computer graphic and cartographic techniques can increase the importance of these methods as research tools in site location studies. Simply by producing maps of individual analytical surfaces, a researcher might gain insights that could be useful in formulating analysis plans or in interpreting analysis results. In addition to traditional maps displaying elevation contours and a hydrology network, maps of new concepts, such as distance to water, aspect, terrain variability, or vegetation diversity, can be produced. Rather than simply producing a map of site locations, the researcher might create a map of an extrapolated site location pattern, which could lead to better insight into the nature of prehistoric land-use patterns. Animation techniques (Moellering 1980) might be used to portray such dynamic processes as landform erosion, air-flow patterns (Tesche and Bergstrom 1978), or changing patterns of settlement through time.

Certain analytical surfaces from a GIS developed for investigating prehistoric patterns of settlement in southeastern Colorado (Kvamme 1984) can be used to illustrate these ideas. Five analytical surfaces from a 230 mi^2 portion of the study area, containing approximately 230,000 cells, each 50 m on a side, are portrayed in Figures 10.11–10.13a. Figure 10.11a is a slope surface. Steep locations (cells) are dark and level locations are light. The surface depicting aspect or principal orientation of the ground surface is shown in Figure 10.11b. In this figure, light shading represents south-facing terrain while dark shading represents north-facing terrain. Note that this surface tends to portray features of the topography related to the drainage systems. The complete hydrologic network is portrayed as the white lines in Figure 10.12a. Also portrayed in this figure are distances to the nearest of these drainages. This information was computed for each of the nondrainage cells, but here, to facilitate display, these data are represented by shading that indicates five categories of distance. A similar map is given in Figure 10.12b, but only a subset of the streams (second Strahler order or greater) is portrayed. Finally, a local relief surface is depicted in Figure 10.13a, which portrays relative terrain roughness and offers contrast between locations of greater and lesser relief. In each cell the range in elevation within a 300 m radius has been determined; high relief values are dark and tend to portray high plateau rim, hilltop, and canyon regions, while low relief values are light and portray plainslike areas. All of these maps portray the same region, but each offers a different way of looking at the landscape.

Perhaps by noting how the distribution of known sites corresponds with these and other surfaces an investigator might better be able to select variables to examine or on which to concentrate in later analyses. Alternatively, an analysis might suggest that certain variables bear a strong relationship with known locations of a particular type of site. In any case, viewing a picture of the mapped variables (Figures 10.11–13a) can give the researcher added insight about his or her findings.

Evaluation of Spatial Statistics

Geographic information systems can be used to examine and evaluate sampling designs and various statistical models. An established regional GIS with known population parameters can be used to investigate (through simulation) the effects of different sampling designs within the region. It might be possible, for example, to investigate a variety of hypothetical sampling designs prior to fieldwork in an effort to fine-tune a particular design to the characteristics of the region under study.

In a similar vein, it is possible to investigate a variety of spatial statistical models and issues. For example, most statistical procedures assume independent observations, but it is usually not possible to meet this assumption when sampling from spatial contexts owing to the presence of positive spatial autocorrelation (Cliff and Ord 1973; also see Chapter 8). Positive spatial autocorrelation has the effect of altering the performance of various statistical models; e.g., levels of significance tend to be overstated (Haggett et al. 1977:329–377). It might be possible to use GIS

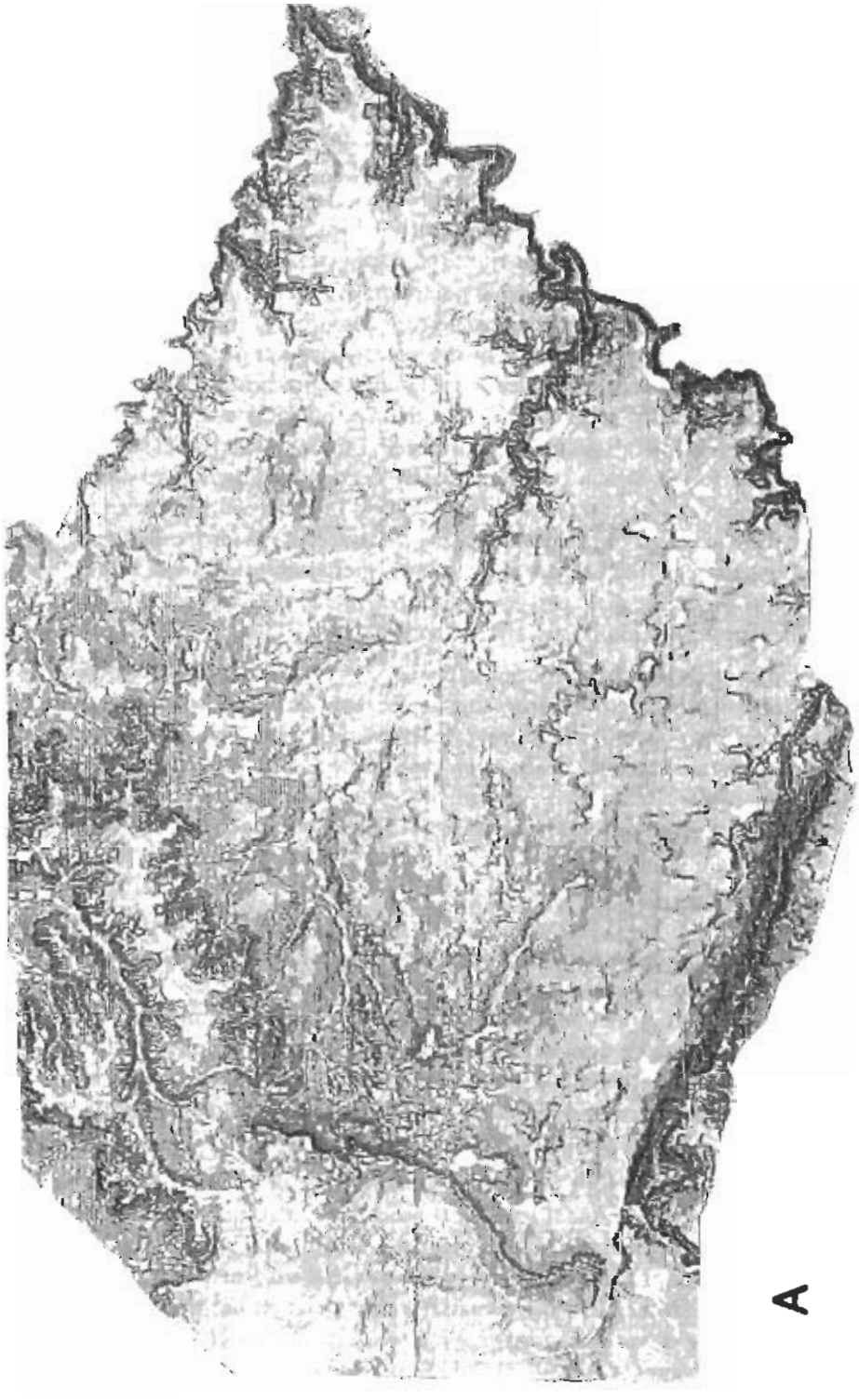
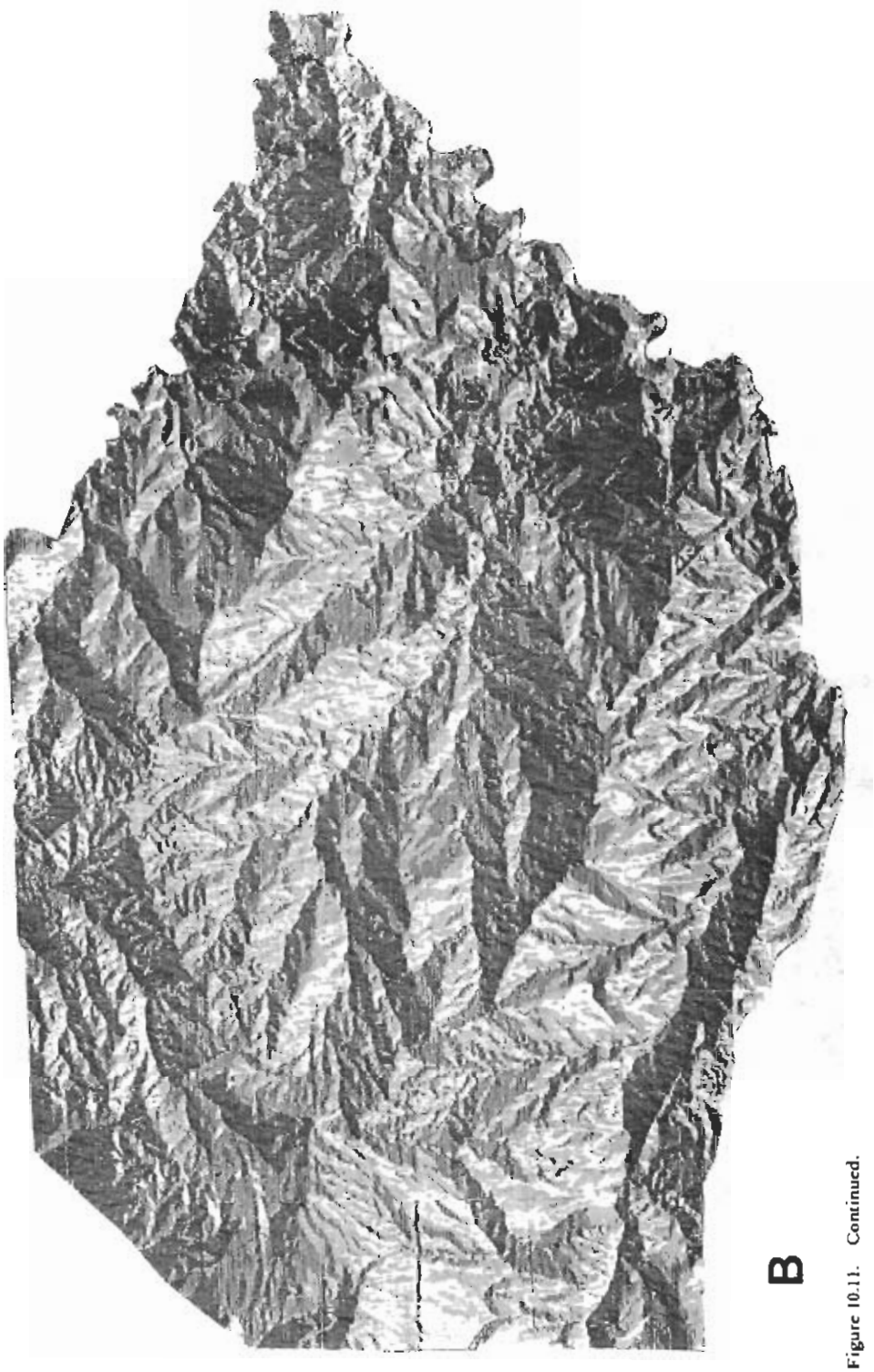


Figure 10.11. GIS-generated surfaces depicting 230,000 individual measurements contained in 50 m² cells across a 230 mi² central Colorado study region. (A) Slope surface: dark regions represent steep ground and light regions represent level ground. (B) Aspect surface: light regions are south-facing and dark regions are north-facing.

(continued)



B

Figure 10.11. Continued.

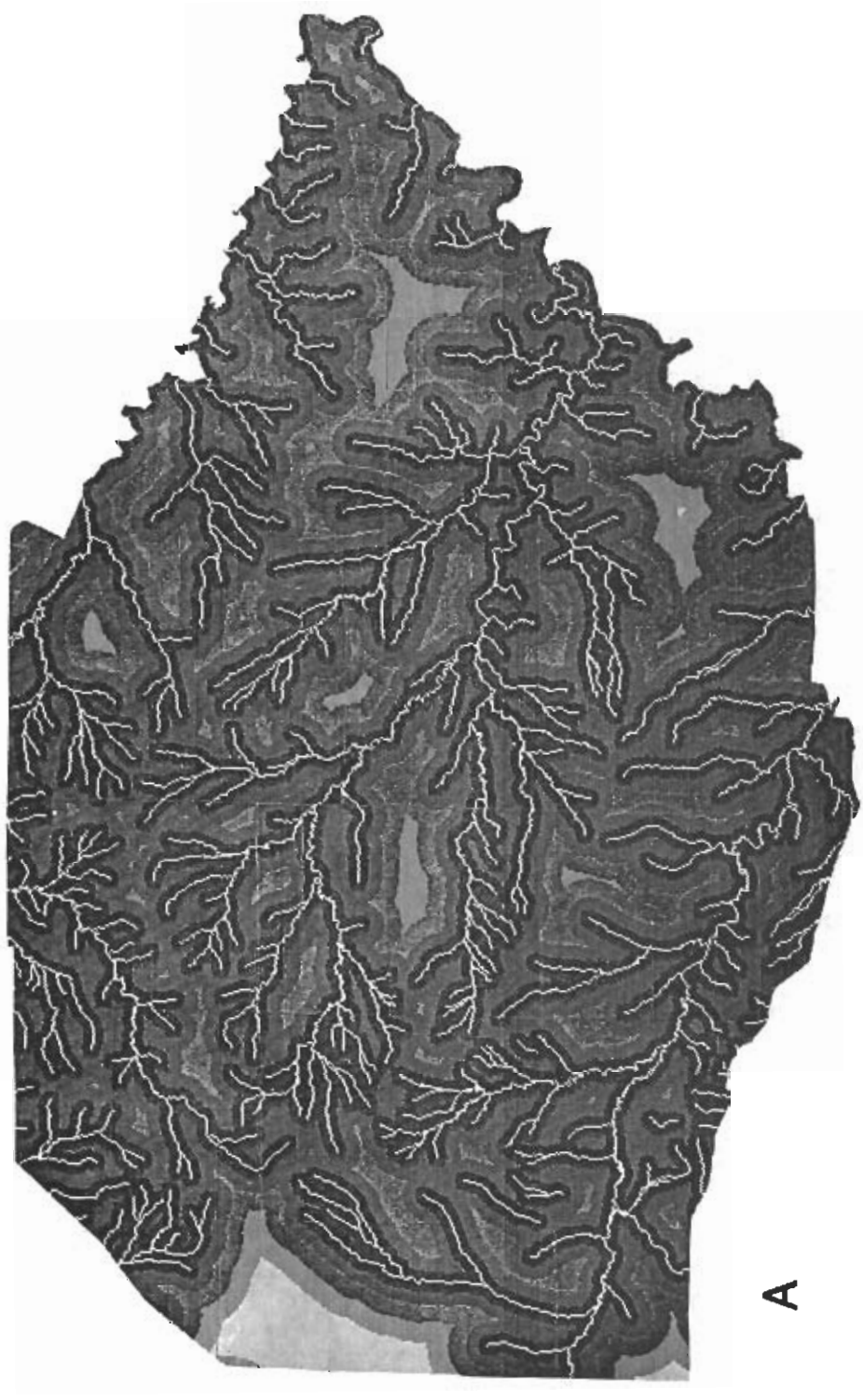


Figure 10.12. GIS-generated surfaces depicting 230,000 individual measurements contained in 50 m² cells across a 230 mi² central Colorado study region. (A) Distances to nearest drainage (shaded) and drainage locations (unshaded). Although a Euclidean distance was calculated for each cell, distances are portrayed in five categories for display purposes. (B) Distances to nearest second Strahler order or greater drainage (shaded) and locations of those drainages (unshaded).

(continued)

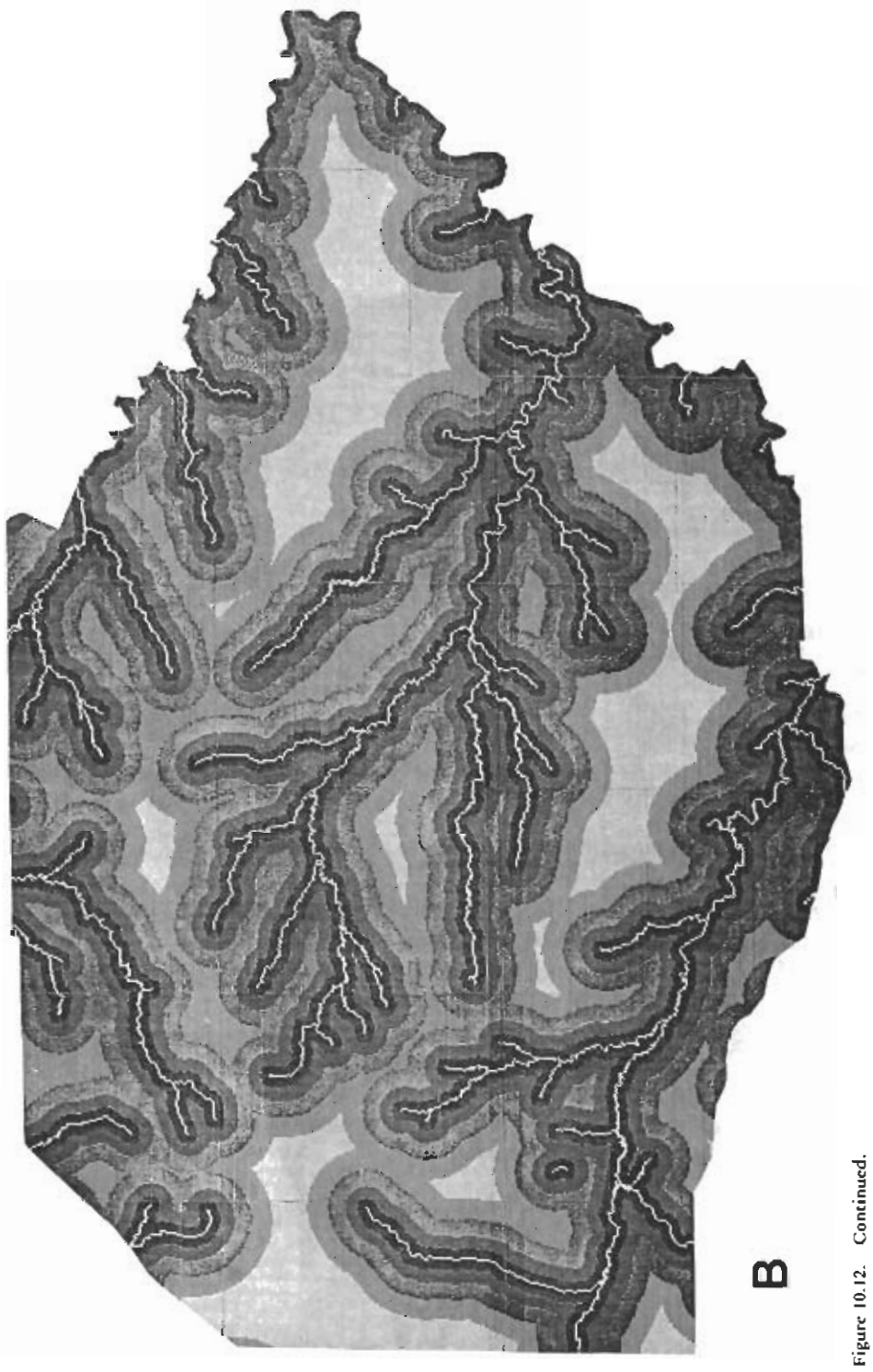
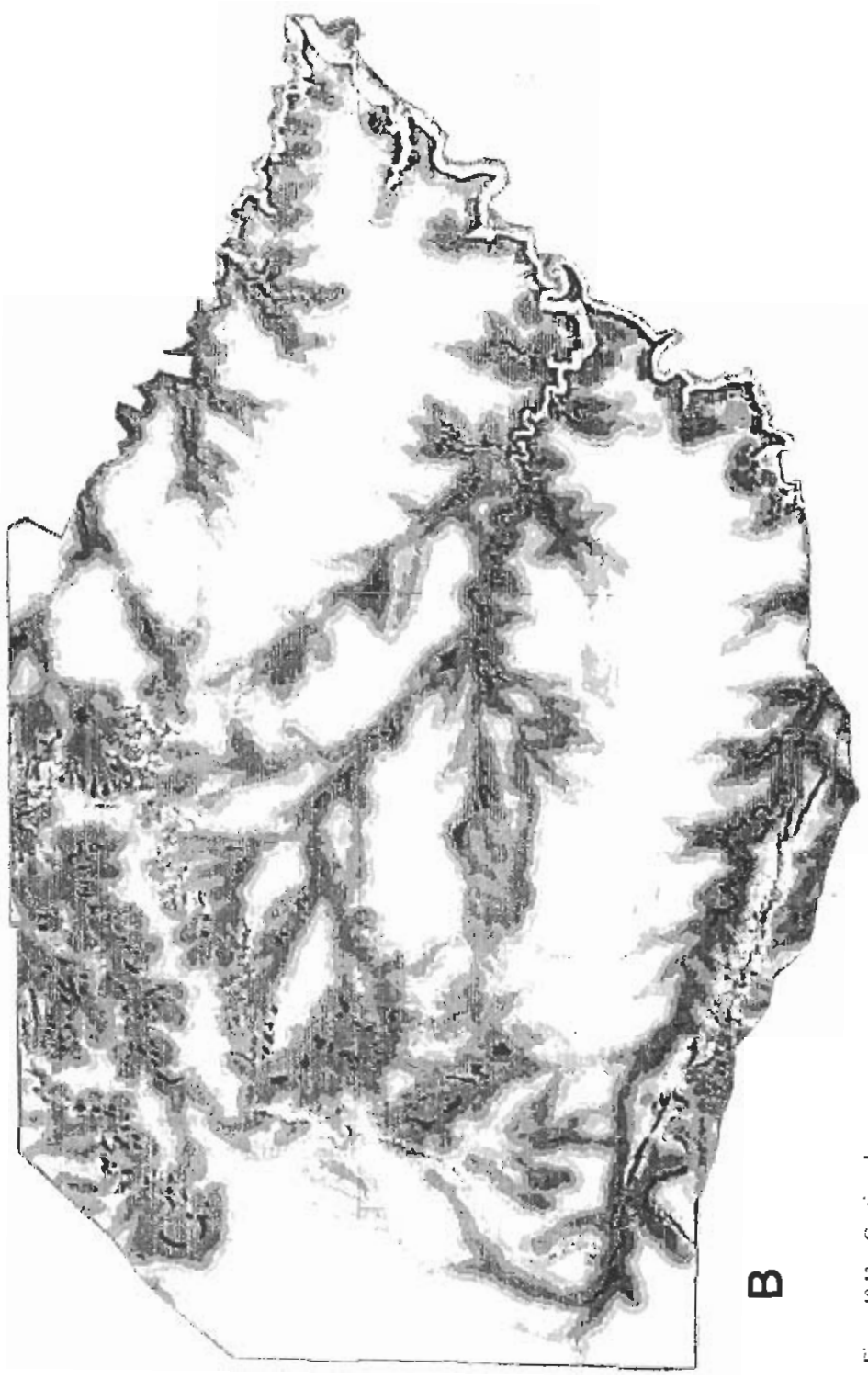


Figure 10.12. Continued.

Figure 10.13. GIS-generated surfaces depicting 230,000 individual measurements contained in 50 m² cells across a 230 mi² central Colorado study region. (A) A local relief surface measured within a 300 m radius of each cell. Dark regions represent high local relief (rough terrain) and light regions represent low local relief (gentle terrain). (B) An archaeological site sensitivity map for open-air lithic scatters. This map presents estimated site-group membership probabilities based on the five surfaces illustrated here (Figures 10.11A-10.13A) and two others. Dark regions represent high site sensitivity (probability) and light regions represent low site sensitivity.

(continued)



B

Figure 10.13. Continued.

data bases as a means of empirically investigating the performance of various statistical models in spatial contexts under controlled conditions, with known autocorrelation structures, perhaps allowing various model corrections to be made (for example, Cliff and Ord 1973).

In a simulation study that used a GIS to investigate levels of spatial autocorrelation under various geographic sampling designs, the effects of this problem with regard to variables commonly used in regional archaeological research were examined (Kvamme 1985). The GIS-based simulation used a 10 by 10 km region as the sampling universe, and for each of five runs of the simulation a different simple random sample and a different regular systematic sample of 100 locations (1 ha grid cells) were selected. Spatial autocorrelation statistics were calculated for each variable for each run. The results indicated extremely high levels of positive spatial autocorrelation regardless of sampling design (some of these results are presented in Chapter 8).

Testing Locational Hypotheses

GIS data bases can be used to test archaeological locational theories and to address other research questions. When a variety of primary and secondarily derived environmental and cultural variables have been previously calculated for a study region in a cell-based GIS, the need for additional measurement can be eliminated. The locations of all known archaeological sites in the region can be easily and rapidly correlated with environmental and other features in the data base. Alternatively, the relationships between GIS data base features and various subsamples of known sites, sites of specific functional types, or sites belonging to a particular period of time can be investigated. For investigators using a control-group approach as a plan for research (see Chapter 8), very large nonsite samples of background environmental or cultural data can be obtained both for model development and for model testing.

Cell-based geographic information systems are ideally suited for an analytical approach to site location research that treats the individual cell (which corresponds to a parcel of land) as the unit of analysis, especially when the cell size is fairly small—e.g., the size of a typical prehistoric site or smaller. Cells that are found to contain artifacts or other cultural remains are simply "flagged" by the computer, thus eliminating site definition problems since the site is no longer the unit of analysis. Relationships between the flagged cells and environmental and other features included in the data base are then examined during model development. Analysis might compare characteristics of cells containing no prehistoric evidence with those of cells that contain prehistoric evidence, for example. Once criteria have been defined for identifying functional site types, site type analyses could be conducted by noting which cells exhibit the required criteria and by flagging cells with a specific site type code. Alternatively, since function is often difficult to determine it might be possible to rank (or continuously measure) cells that contain cultural evidence according to artifact diversity or to amounts of inferred prehis-

toric activity, using various threshold levels of amounts of prehistoric evidence. Various location models might then be developed in which the dependent variable is an index of artifact counts, diversity, or levels of prehistoric use.

GIS data bases are well suited for testing certain types of site locational theories. It might be postulated, for example, that certain kinds of archaeological sites in a study region should be located close to sources of water. A GIS data base could be used to determine empirical distances to water at known sites of the type under investigation in order to test this hypothesis. It should be recognized, however, that *all* parts of the study region might generally lie close to water sources. Hence, even if the sites tend to be located close to water sources, this tendency could be a result of the nature of the background environment rather than of prehistoric selectivity, for example. Measurements from the background environment might yield a distribution of distances to water identical to that for sites, which would suggest no selectivity, or the distributions might be radically different, suggesting selectivity. GIS techniques are ideal for investigating such an issue because they can provide many thousands of background measurements of environment against which known site distributions can be compared.

To illustrate the power of geographic information systems for analysis purposes, a simple histogram is presented in Figure 10.14a that illustrates the Euclidean distance to the nearest drainage of Strahler order rank two or greater as measured by a GIS in 230,000 contiguous cells (50 m on a side) in central Colorado (Kvamme 1984). This figure clearly illustrates the nature of the background environment in this region with respect to this variable. The histogram of the same variable measured only at the locations (cells) of nearly 600 known open-air lithic scatters in the area portrays a distinct tendency for the sites to be located in relatively greater proximity to second-order streams (Figure 10.14b). For example, 50 percent of the sites occur within 150 m of second-order or greater drainages, while only 17 percent of the study region as a whole exhibits a similar proximity to these drainages; 90 percent of the sites lies within 950 m of the drainages, while only 69 percent of the study region lies within this distance. Since the sample of open-air sites was obtained through a random sampling design, the patterning apparent in Figure 10.14 is difficult to refute and points to the tremendous potential of geographic information systems for archaeological locational investigations.

Locational Modeling

GIS techniques are well suited for the development, testing, and application of archaeological locational models of any type (see, for example, Chapter 8). The only limitations are that appropriate forms of geographically distributed information (including remotely sensed data and specialized map or aerial photograph data) must be merged into the data base and that the cell resolution or size must be appropriate for the modeling problem.

In developing quantitative models based on probability or mathematical functions of multiple geographic variables, geographic information systems can be

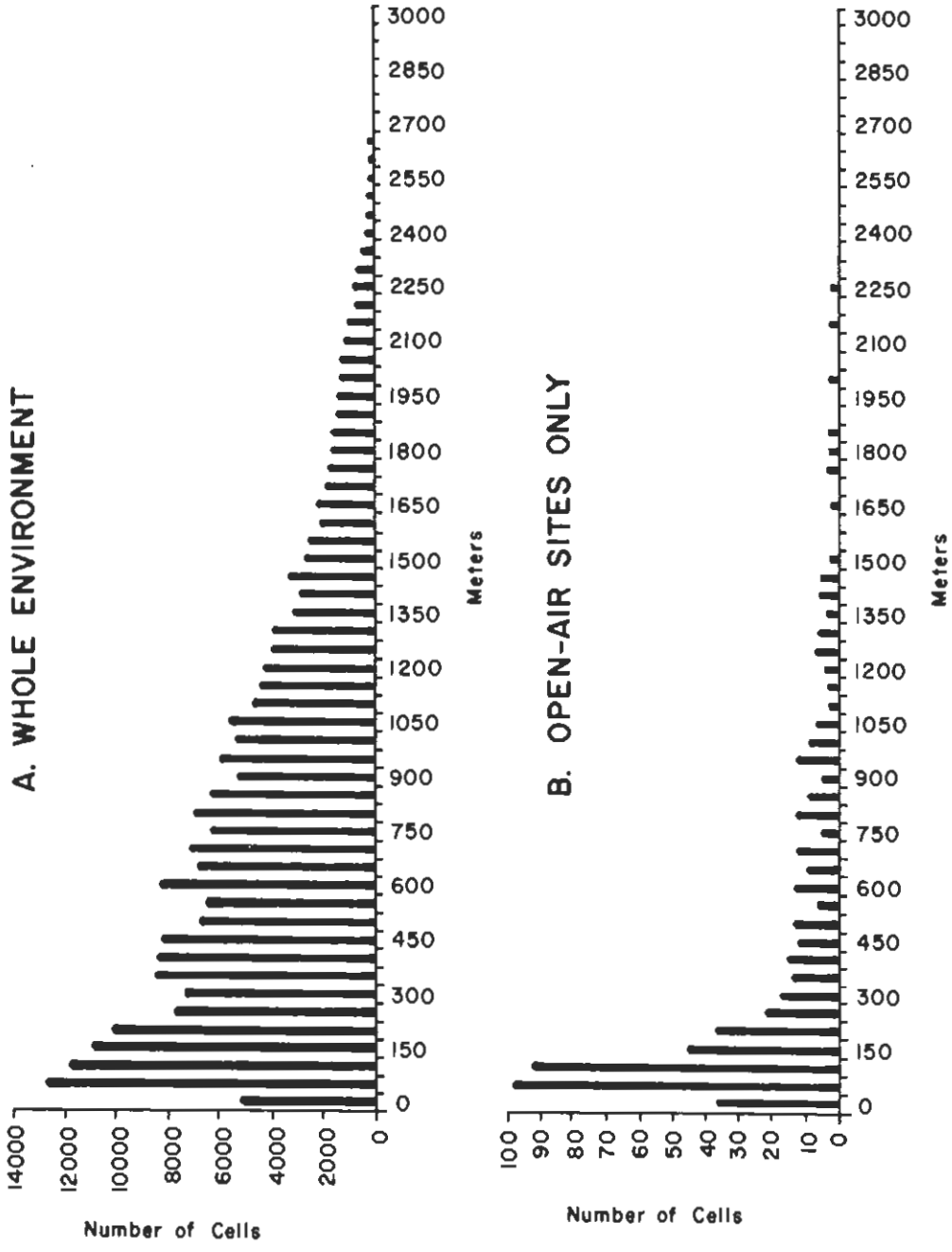


Figure 10.14. Histograms of Euclidean distances to the nearest second Strahler order or greater stream in a central Colorado study area. (A) 230,000 distances measured every 50 m across the study region represent the nature of the environment as a whole with respect to this variable. (B) 538 distances measured only at the locations of a representative sample of open-air lithic scatters. A comparison of the two histograms suggests a tendency for sites to be located in proximity to these drainages.

used to obtain environmental and other variables at the locations of known sites (or site types) to provide the basic analysis data. During model testing, geographical data merged with a second sample of sites can be retrieved and used as a basis for a variety of accuracy tests (see Chapter 8). Finally, geographic information systems can be employed to specify the results of a model across a region of study by applying the model (i.e., the probability or mathematical function) to the data stored in each cell and producing a map of the results.

Figure 10.13b illustrates a prehistoric "site probability surface" derived through a logistic regression technique (see Chapter 8) for the site class of "open-air lithic scatters" in the central Colorado project described earlier. This model is based on a sample of nearly 300 known open-air sites and a control group of approximately 1200 locations representing the background environment (nonsites). In each of the 230,000 cells in this figure an estimated probability of site-class membership was derived, conditional on seven environmental variables within the GIS data base (including those illustrated in Figures 10.11–10.13a). Computer cartographic techniques were used in Figure 10.13b to shade cells having p-values nearest to 1 with dark tones, to shade cells with p-values near 0 in light tones (or unshaded), and to shade cells with intermediate p-values in intermediate tones. The result is a visual representation of the extrapolated pattern of open-air site placement, based on the sample data.

This model was also tested using a GIS. Test results from an independent validation sample of an additional 300 open-air sites and 1200 background locations suggest that about 95 percent of the sites (92–97 percent at an approximate 95 percent level of confidence) should occur in all the shaded zones of the map, although these shaded areas constitute only 62 percent of the total land area. The results also indicate that approximately 20 percent of the sites (16–25 percent at ca. 95 percent confidence levels) should occur in the highest sensitivity zone (the darkest shading level), which covers less than 4 percent of the total land area (Figure 10.13b).

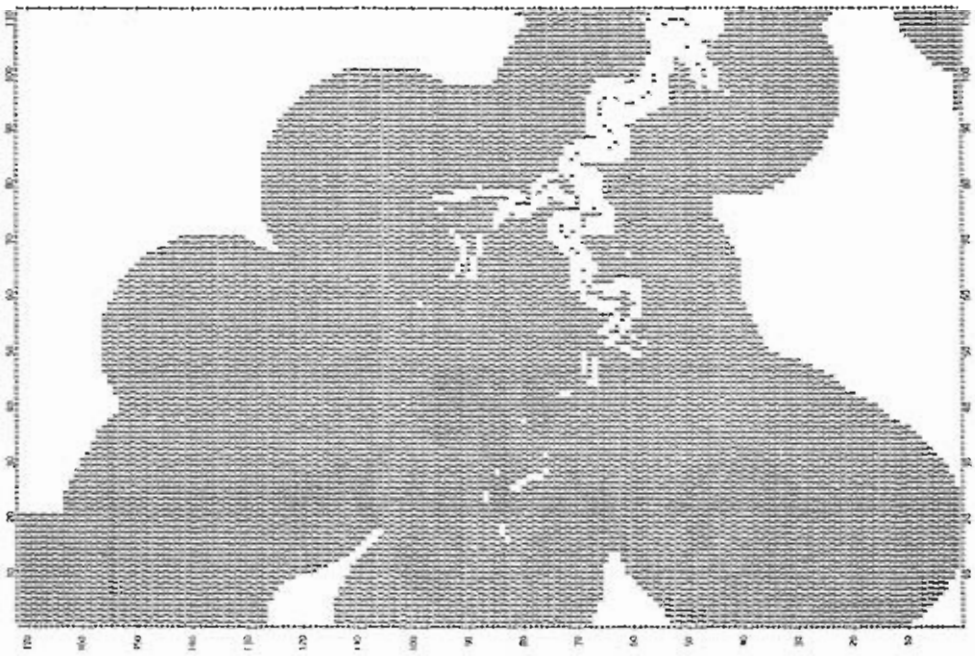
For deductively derived modeling approaches, model development cannot be carried out within a GIS framework since these approaches do not begin by seeking patterns in empirical data. Such models are based on theoretical principles concerning human choice and settlement behavior and consist of deductions about the locations at which human occupation should occur. Once these models have been established, however, geographic infotmation systems can be used for model testing and broad-area applications.

One problem in applying many deductively based models lies in data requirements. For example, to apply central-place modeling techniques (Johnson 1977), which assert the importance of central places to a regional pattern of settlement, one must know the locations of contemporary central places. Gravity models (Hodder and Orton 1976:187), which emphasize the importance of specific natural resources (e.g., food resources or lithic quarries) or cultural entities (e.g., road networks or central places), require locational data for each of these phenomena. Models based on caloric cost-benefit or energy calculations (e.g., Casteel 1972;

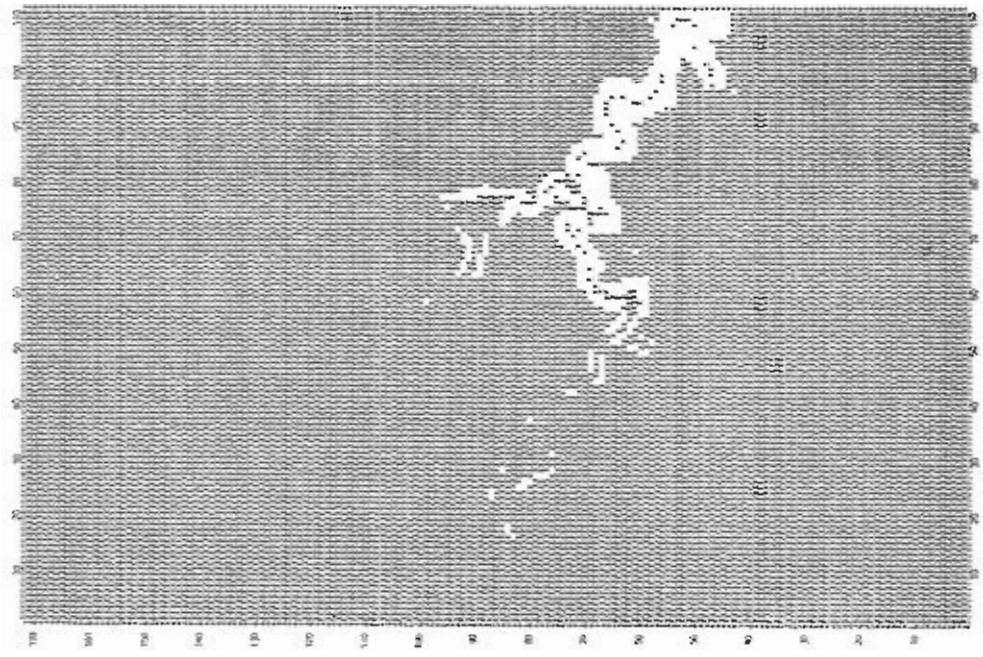
Zubrow 1971) require detailed environmental information. In one modeling approach based primarily on environmental data, Jochim (1976) was able to arrive at several deductions concerning hunter-gatherer settlement by synthesizing a wide range of ethnographic and other information. Unfortunately, the required data for application of the model, which included detailed information about such items as the food potential of several prehistoric plant and animal species, their relative proportions, and their seasonal abundance, were so difficult to obtain in a reliable form in the time period and region to which the model was applied (the Mesolithic of southern Germany) that it was difficult to realize the full potential of the model.

GIS techniques may offer a solution to some of these problems, provided that the relevant data can be gathered and incorporated within a GIS framework. A variety of map sources or even zoological models might be used, for example, to describe the distributions of certain species of interest, and remote sensing techniques might be used to identify prehistoric central places, road networks, major plant groups, favorable plant diversity, or other features. Once the archaeological locational model is formulated and made operational in computer terms, computer mapping techniques in conjunction with GIS features provide an easy means of applying the model across the region of interest. Testing of any model demands similar procedures regardless of how the model is developed (testing procedures are described in detail in Chapter 8), and as described above, geographic information systems are well suited for model testing purposes.

The test study region of 19,000 grid cells that was used to illustrate the quantitative models in Chapter 8 can be used to indicate the potential of geographical information systems in an a priori model specification perspective. Whether an archaeological locational model is derived simply through a series of "shotgun" questions put to a GIS or through a series of deductions concerning the interrelationships between certain environmental features and the positioning of human settlements in space, a GIS can be used to map the results of the modeling process. As a simple example, a base model might specify that settlements should occur on ground surfaces with slopes less than or equal to a 12 percent grade (Figure 10.15a). The next refinement of this model might then suggest that settlements should be found within a fixed distance, say 1000 m, of relatively secure water, such as second Strahler order or greater streams (Figure 10.15b). Finally, the model might be amended to include the requirement that particular settlement locations (e.g., those of winter villages) will have a south-facing orientation (Figure 10.15c). At each stage in the development of this model, accuracy, in terms of the percentage of known sites correctly classified and the percentage of the region classified by the model as "site-present," could be assessed by the GIS, providing ongoing and interactive model performance indications.

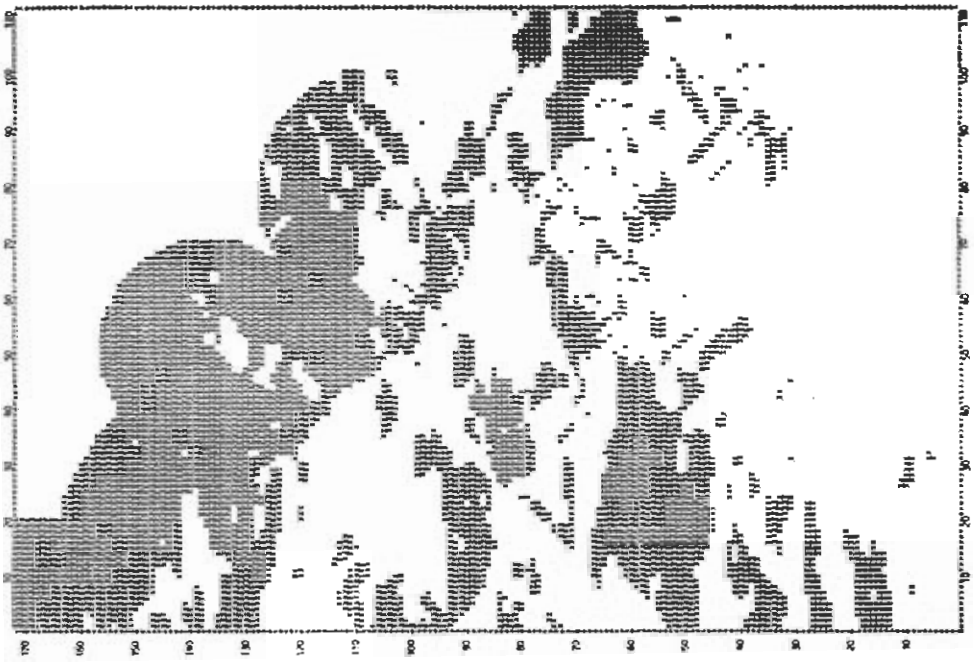


B



A

Figure 10.15. Illustration of successive and cumulative selections made by a GIS. (A) All locations with a ground slope less than or equal to a 12 percent grade. (B) All locations within 1000 m of secured water (second Strahler order or greater streams). (C) All locations with a south-facing orientation. *(continued)*



C

Figure 10.15. Continued.

BRIEF OVERVIEW OF THREE COMMON GEOGRAPHIC INFORMATION SYSTEMS

Timothy A. Kohler

MOSS/MAPS

MOSS (Map Overlay and Statistical System) is a GIS originally developed by the Western Energy and Land Use Team, U.S. Fish and Wildlife Service (Table 10.1, above). It has been in continual development over the past few years with cooperation from the Bureau of Indian Affairs, the Bureau of Land Management, the Forest Service, the Geological Survey, and the Soil Conservation Service (Lee et al. 1984). Thus, unlike most geographic systems it is in the public domain, although a superset of MOSS is marketed by Autometric of Fort Collins, Colorado, a firm that is also developing a more advanced GIS, based on MOSS, called DELTAMAP (Reed 1986). Most storage and processing in MOSS is in a vector or polygon format, although some raster capabilities are available.

Additional raster capabilities, designed in part to allow the incorporation of data derived from digitized images, are available through the Map Analysis and Processing System (MAPS) subsystem, an extensively enhanced version of MAP, originally developed at Yale University. To some extent, MAPS and MOSS can pass files back and forth. Input to MOSS is through MAPS; AMS, the Analytical Mapping System; or ADS, the Automated Digitizing System. Enhanced cartographic plotting, beyond the normal capabilities of MOSS or MAPS, is provided by the Cartographic Output System (COS).

Beyond the general capabilities of geographic information systems as described earlier in this chaptet, MOSS and MAPS have special capabilities that are of interest for predictive locational modeling of archaeological sites. These include routines that

— collect a random sample of points, lines, or polygons for further analysis or for input to statistical procedures

— measure the distance between any two points along a path (which need not be straight) or along a straight line

— determine the length of all lines of each subject (e.g., first-order streams) in a line map or the total distance around each subject in a polygon map

— identify locations within a specifiable distance of a point, line, or polygon subject type

— produce a three-dimensional display of any integer-valued continuous map

— create a map of azimuthal aspect or a slope map from a continuous elevation map

— create a map showing the visibility of locations from a specifiable observation point or points

— create a cross-sectional image between any two points (usually this routine is used for elevation data, but it is suitable for any continuous map)

— create a map showing the minimum effort path to a target cell; the analyst can assign weights to various features acting as partial barriers in the path-finding process (an example of a fairly common GIS capability for *corridor analysis*)

— create a map showing the steepest downhill path through varying terrain (essentially the path along which water would flow) from a target area

The MOSS/MAPS package provides very flexible routines for overlay and neighborhood analysis, map description, and data management. A principal advantage of this package is that it is used and supported by numerous federal agencies, so that new features are being added to it at a rapid rate. At present MOSS/MAPS has only very limited capabilities for inferential statistical analysis (Table 10.3). Versions are available for 16- and 32-bit microcomputers, minicomputers, and mainframes.

TABLE 10.3.

Statistical functions (beyond simple descriptive statistics) available in three commonly utilized systems

Function	*MOSS/MAPS*	*IDIMS*	*VICAR/IBIS*
Supervised cluster analysis	—	X	X
Unsupervised cluster analysis	—	X	X
Principal components analysis	—	X	X
Least squares analysis	—	—	X
Divergence calculations	—	X	—
Cross-tabulation	X	—	—

IDIMS

Unlike the public-domain system MOSS/MAPS, the Interactive Digital Image Information System is a commercial product of the Electromagnetic Systems Laboratory, Inc., in Sunnyvale, California. Like VICAR, which is discussed below, IDIMS is primarily an image-processing system; for this reason, data are organized in a raster format, and many functions that address problems specific to the processing of digital images, such as image-enhancement routines, are available. Many other IDIMS functions are useful for more general kinds of spatial analysis, however, so it also warrants consideration as a GIS. IDIMS incorporates a data-entry component, the Geographic Entry System (GES), and the Earth Resources Inventory System (ERIS) for data base management and statistical functions (Electromagnetic Systems Laboratory n.d.; Hansen 1983).

Special features beyond functions that are routine in an image-processing or geographic information system, or that might be of special interest for archaeological locations analysis, include

— a procedure for overlaying up to 10 maps (or images) at one time, rather than the two at a time possible in MOSS/MAPS

— a procedure that passes a three-by-three-cell moving window across a land-cover map to create a diversity index

— procedures for creating slope and aspect maps from digital elevation data

— a procedure for creating a shaded relief map from a digital elevation map with the sun in a specifiable location

— a procedure for creating a proximal map that assigns each cell to the nearest given *x,y* location

— various procedures for generating random samples of images for further analysis

IDIMS runs on a minicomputer and is used by several large federal agencies. Hansen (1983) has discussed the creation of a "generic" GIS through combining the most useful features of MOSS, MAPS, IDIMS, and their various allied programs for data entry, management, and display.

VICAR/IBIS

The Video Image Communication and Retrieval (VICAR) system was developed at the Jet Propulsion Laboratory (JPL) to process image data from the planetary exploration programs of the late 1960s and 1970s (Bracken et al. 1983; Hart and Wherry 1984). Unlike MOSS/MAPS and IDIMS, VICAR is designed to run on large-scale digital computers and is normally restricted to IBM systems, since a substantial proportion of its code is written in IBM 360/370 Assembler Language. A subset of VICAR, called mini-VICAR, was developed to run on DEC minicomputers, but it appears that this system is no longer actively used. A DEC VAX version of full VICAR is now in use at JPL, however. Like MOSS/MAPS, VICAR is in the public domain. With well over 100 application programs running at about 25 installations around the world, VICAR is a very powerful and widely utilized image-processing system.

The Image Based Information System (IBIS) is an enhancement to VICAR also developed at JPL (Bracken et al. 1983; Zobrist and Bryant 1979). The IBIS programs give VICAR/IBIS some of the capabilities of a GIS, including overlay analysis and vector-to-raster conversion, which allows geo-coded information not normally available in raster (cell) formats (such as maps) to be analyzed.

The majority of VICAR application programs are specialized for image processing, a task that may sometimes be important in predictive archaeological modeling—particularly when map-based data are unavailable, It is also important

to remember that most modern, map-based data are the result of image interpretation of some sort. A few of the VICAR programs of potential importance to locational modeling include

— a multivariate classifier program using Bayes's maximum likelihood algorithm, which yields a classified image (map) and, optionally, a confidence map for that classification. This program accepts input either from a *supervised* classification analysis, in which the user specifies certain "training areas" on which the classification function is to be based, or from an unsupervised cluster analysis

— a multivariate classifier program that uses a combination of parallelepiped and maximum likelihood techniques, accepting input from either a supervised or an unsupervised analysis

— a program for performing edge enhancements and, optionally, for making edge existence decisions

— a principal components analysis of up to 12 input variables

— a least-squares program that will, among other things, calculate and display trend surfaces and residuals from trend surfaces

— a program that simulates the effect of shading from a specifiable angle of illumination on any continuous image

The fact that VICAR/IBIS typically runs on large mainframe computers has both advantageous and disadvantageous aspects. In installations with which I am familiar, VICAR/IBIS runs as a "batch" program, meaning that jobs are submitted, and the output later (possibly much later) received, with no intermediate interaction between the user and the processing system. Obviously, it is desirable to have fast response to user query in an interactive mode, as is typically the case for geographic information systems running on mini- or microcomputers. There is great analytical utility in being able to see the mapping of some function unfold before your eyes, perhaps to be interrupted and modified if necessary in its early stages. On the other hand, some batch systems, such as VICAR/IBIS, have a huge variety of sophisticated functions, and their mainframe implementation allows the use of very large data bases. As new data storage technologies, such as laser disks, become available, and as the cost of data storage continues to drop, one of the advantages of mainframe-based systems will disappear. On the other hand, as cheaper and more powerful local workstations begin to share processing with mainframes, the easy dichotomy between mainframe- and microcomputer-based geographic information systems will also become fuzzy, and batch systems will probably become things of the past.

MODEL DISPLAY VS MODEL BUILDING

Timothy A. Kohler

Earlier in this chapter, Kvamme discussed many realized or potential applications of geographic information systems to general spatial research in archaeology, including the construction, testing, and use of predictive locational models. Within the category of model building, a distinction can be made between the processing necessary to build a data set suitable for inferential statistical testing and the application of inferential procedures (for example, linear regression) to discover the "best" locational model. It is important for managers to realize that, in their present phase of development, most geographic information systems are much better suited to the first task than to the second. Constructing an inferential model of site location inevitably involves the application of inferential statistics to surveyed areas that contain or are devoid of archaeological resources. Geographic information systems give unsurpassed power for the extrapolation of such models—once constructed—to the area from which the samples were originally drawn, but actual inferential statistical functions available in many geographic information systems are rather limited (Table 10.3). This is not a fatal weakness for the application of a GIS for model building if the GIS has the ability to format a file for use by a general-purpose statistical package, such as SAS or SPSS, as is usually the case. It does mean, however, that a GIS is usually not the only software needed for the analysis of spatial data.

Of course, geographic information systems are an important aid in model construction since they can be used to *collect* data to be passed to an inferential statistical analysis. As pointed out by Kvamme, anyone who has conducted a quantitative settlement pattern analysis—examining the distances from known archaeological resources and random points to various features of the natural environment and evaluating the composition of catchments around both sites and random points—knows how tedious and prone to error these hand measurements can be. In a GIS suitable for archaeological analysis, such measurements can be made automatically for any of the available data planes or maps. These measurements constitute secondary surfaces that can be stored as new maps on which the locations of sites and points without sites are replaced by measurements of catchment composition and distances to critical resources around these points. These measurements can then be passed to another system for statistical analysis, and in this manner the most tedious portion of model construction has been automated. Perhaps when it is easier to consider variables related to catchment values and distances to resources, these variables will be used more frequently and effectively in predictive locational modeling than they have been to date.

IMAGINARY SESSION WITH A GENERIC GIS

Timothy A. Kohler

Despite the growing literature about and increasing accessibility to geographic information systems, these systems remain mysterious to most archaeologists. What follows is a poor man's substitute for the only experience that can really convey both the usefulness and the limitations of these systems—a "hands-on" session. This example illustrates how a GIS might be used to map a model that has *already* been developed, either by using the GIS for data collection or by some other means. Limiting this example to mapping rather than development of a model should help the reader who has no acquaintance whatsoever with geographic information systems to understand how they work. Additionally, as pointed out in the previous section, building an inferential model is essentially a statistical task in which the GIS serves as a technical assistant for data collection and management. The specific techniques discussed are more appropriate for image-processing-based systems (such as IDIMS) than for many geographic information systems, and there would certainly be more efficient ways to approach this task on some systems.

You sit in front of a high-resolution graphics terminal attached to a minicomputer or a "supermicro" running a relatively advanced GIS. The most tedious and expensive work—digitizing various maps for the data base, correcting digitizing errors, geometrically correcting remote sensing imagery, tying that imagery into ground control points, and so forth—has already been done. Previous researchers have interpreted available Landsat imagery to yield digitized maps of vegetation type and density and of current land use. Likewise, digital elevation models available on computer tape from the USGS (Elassal and Caruso 1983) providing elevations for points at 30 m intervals have already been processed to yield secondary maps of slope and aspect. Each of these digitized maps has been stored on disk or tape and is accessible to the computer, and each constitutes a *data plane* or *theme*. Themes available to you for our imaginary GIS session are shown on Table 10.4. These data are available for an area about 51 km on a side, the largest area your monitor can display at a resolution (picture element, pixel, or cell size) of 50 m on a side. More than a million (1024^2) pixels are displayed on your screen, which shows an area equivalent to that portrayed by about 16.5 USGS 7.5-minute topographic quadrangles. You can enlarge any portion of this area to fill the whole screen if you wish to see a subset of the area in more detail.

Relatively low quality copies of the contents of the screen can be obtained quickly and cheaply in black-and-white on a peripheral dot-matrix printer; high-quality color copies can be obtained using a peripheral pen plotter or a high-resolution color ink-jet printer. The system at your disposal cost somewhere between $40,000 and $125,000 and so must be shared by many different users, most of whom are involved in natural resources inventory and analysis.

TABLE 10.4.

Data themes available for your GIS session

Description	*Organization*
Archaeological sites with attributes for type and age	Polygon
Aspect	Cell
Elevation	Cell
Extent of archaeological survey	Polygon
Modern land use	Polygon
Roads/streams with attributes for types/orders	Line
Soil type	Polygon
Slope	Cell
Vegetation types	Polygon
Vegetation density	Polygon

You wish to map a simple site location model that predicts, for example, that frequencies for two types of sites will be relatively high in locations satisfying two slightly different sets of criteria. Requirements for the first site type are locations with

- less than 5° slope,
- no more than 8000 ft elevation, and
- permanent water and piñon-juniper woodland no farther than 0.25 km away.

The second class of sites is likely to occur in areas with

- no more than 10° slope,
- seasonal or permanent water no more than 0.25 km distant,
- no more than 7500 ft and no less than 6000 ft elevations,
- arable soils no more than 0.1 km distant, and
- locations at the base of a slope.

You wish to create one map showing those areas most likely to have site types 1, 2, both, or neither.

There are many ways to approach this problem; details of the "best" approach depend on the characteristics of the particular GIS at your disposal. One likely approach—ignoring technical details and considering only the general strategy—would be to select all locations for each site type on each data plane that are favorable to settlement and code them with a 1, coding all other areas with a 0. Once this operation is completed for each relevant data plane (that is, for each map of a particular variable or environmental characteristic), the four data planes (in the case of site type 1) or five data planes (in the case of site type 2) can be electronically overlaid, with values from the same location on each map being summed together. This procedure is analogous to overlaying a series of accurately positioned and extremely detailed mylar maps to produce one new map in which each location is

the sum (or some other function) for that location of the information presented on all the overlain maps. The next step would be to recode all areas that yielded a sum of 4 in the first analysis to 1, with other areas assigned 0; all areas with a 5 in the second analysis would be recoded to a 2, with other areas assigned 0. In this fashion two summary maps would be created, one for each site type. These, in turn, would be overlaid to create a final map in which any location with a 1 would meet the criteria for site type 1 only; any location with a 2 would meet only the criteria for site type 2; and locations with a 3 would meet the requirements of either type. Locations coded 0 would be considered not to meet the requirements of either type.

With one exception, the processing to be done *within* any data plane prior to overlaying the separate data planes is simple and straightforward. For example, the process of selection according to a range of elevation and slope criteria relies on a very basic ability of geographic information systems to reclassify or renumber data planes. In the analysis for site type 2, for example, a new copy of the master slope map would be made in which all locations with a slope of 10° or less would be coded 1, while other locations would take on a value of 0.

Other basic GIS capabilities are illustrated by the operation of finding locations within a certain radius of some environmental feature or attribute (such as within 0.25 km of seasonal or permanent water). One way to do this is to pass a "moving window" with a radius equal to the maximum distance allowable from the feature across the pixels that constitute the "electronic landscape." Any point within the allowable distance could be flagged on a new map with a certain value, perhaps a 1, while other locations would take on a value of 0. Another method, which is usually more efficient, employs a function that expands the perimeter of the feature of interest by the proper distance. These functions create a concentric zone of specifiable width around a point, line, or polygon, an operation that is frequently useful in archaeological spatial analysis. One can, for example, specify a vegetation zone (piñon-juniper) to be used as a target; the width of the concentric zone to be created around any occurrence of this vegetation type; and the numbers to be assigned to locations within this expanded zone. In the example discussed above, this expansion function would be employed twice during the mapping of possible site type 1 locations—once on the roads/streams data plane, using permanent streams as a target, and once on the vegetation data plane, using piñon-juniper as a target.

The one exception mentioned above to the rule of relatively simple information processing within each data plane involves the problem of finding locations near the base of a slope. In most geographic information systems this would require a several-stage process (more complex than we need to describe here) that might, when completed, give less than perfect results. This example is included to demonstrate that not all results that are easy for a human to achieve (as locating areas at the base of a slope might be) are necessarily easy to achieve via a computer, given current technology.

The entire analysis just described might take a couple of hours with a large computer or a couple of days with a smaller one. In either case, the great time and

expense incurred in collecting and digitizing the data, once completed, need not be repeated, and users with different goals can profit from the accumulated, organized, and highly accessible information in the GIS. Even with a smaller, slower computer, results are achieved much more rapidly and accurately than if the work was done by hand, assuming that the data base is in place.

Ken Kvamme reiterates his acknowledgment of those persons and institutions mentioned at the end of Chapter 8. Tim Kohler would like to thank Judy Hart, David Wherry, and Robert Wright for comments on an earlier version of his portion of this chapter.

REFERENCES CITED

Allan, M. M.
1978 DTM Application in Topographic Mapping. *Photogrammetric Engineering and Remote Sensing* 44:1513–1520.

American Farmland Trust
1985 Survey of Geographic Information Systems for Natural Resources Decision Making at the Local Level. Copies available from American Farmland Trust, Washington, D.C.

Bracken, Peter A., William B. Green, and Robert M. Harilick
1983 Remote Sensing Software Systems. In *Manual of Remote Sensing*, vol. 1, 2nd ed., edited by R. N. Colwell, pp. 807–839. American Society of Photogrammetry, Falls Church, Virginia.

Brown, Patricia Eyring, and Connie Lynn Stone
1982 *Granite Reef: A Study in Desert Archaeology*. Arizona State University Anthropological Research Papers 28, Anthropological Field Studies 3. Tempe.

Bryant, Nevin A., and A. L. Zobrist
1977 Integration of Socioeconomic Data and Remotely Sensed Imagery for Land Use Applications. In *Mapping with Remotely Sensed Data: Proceedings of the Second Annual William T. Pecora Memorial Symposium*, pp. 120-130. American Society of Photogrammetry, Falls Church, Virginia.

Calcomp
1983 *Calcomp 9000 Series Digitizer Operator's Manual 500090-1*, revision B. Talos Systems, Scottsdale, Arizona.

Casteel, Richard W.
1972 Two Static Maximum Population Density Models for Hunter-Gatherers: A First Approximation. *World Archaeology* 4(1):19–39.

Cliff, A. D., and J. K. Ord
1973 *Spatial Autocorrelation*. Pion, London.

Collins, Stanley H., and George C. Moon
1981 Algorithms for Dense Digital Terrain Models. *Photogrammetric Engineering and Remote Sensing* 47:71–76.

Cordell, Linda S., and Dee F. Green, editors
1983 *Theory and Model Building: Refining Survey Strategies for Locating Prehistoric Heritage Resources.* Cultural Resources Document No. 3. USDA Forest Service, Albuquerque.

Creamer, Winifred
1985 The Upper Klethla Valley Survey: Computer Generated Maps of Site Location. Paper presented at the 50th Annual Meeting of the Society for American Archaeology, Denver.

Delfiner, P., and J. P. Delhomme
1975 Optimum Interpolation by Kriging. In *Display and Analysis of Spatial Data*, edited by John C. Davis and Michael J. McCullagh, pp. 96–114. John Wiley, New York.

Doyle, Frederick J.
1978 Digital Terrain Models: An Overview. *Photogrammetric Engineering and Remote Sensing* 44:1481–1485.

Edwards, Kathleen, and R. M. Batson
1980 Preparation and Presentation of Digital Maps in Raster Formar. *The American Cartographer* 7:39–49.

Elassal, Atef A., and Vincent M. Caruso
1983 *Digital Elevation Models.* U.S Geological Survey Circular 895-B. User Services Section, NCIC/USGS, Reston, Virginia.

Electromagnetic Systems Laboratory (ESL)
n.d. IDIMS Functional Guide. Electromagnetic Systems Laboratory, Sunnyvale, California. On file, EROS Data Center, Sioux Falls, South Dakota.

Ericson, Jonathon E., and Raymond Goldstein
1981 Workspace: A New Approach to the Analysis of Energy Expenditure within Site Catchments. *Anthropology UCLA* 10:21–30.

Erikson, W. K., L. B. Hoffman, and W. E. Donovan
1983 The Microcomputer Workstation: An Alternative Hardware Architecture for Remotely Sensed Image Analysis. Paper presented at the Eighth William T. Pecora Memorial Remote Sensing Symposium, Sioux Falls, South Dakota.

Estes, John E., Stephen Guptill, and Jeffrey L. Star
1984 Digital Geographic Information Systems. Readings prepared for Course No. 17, School of Engineering and Applied Science. George Washington University, Washingron, D.C.

Green, Dee F., and J. B. Stewart
1983 Computer Generated Research Aids Using the Forest Service DTIS II System. *Advances in Computer Archaeology* 1:4–25.

Haggett, P., A. D. Cliff, and A. Frey
1977 *Locational Methods*, 2nd ed. John Wiley, New York.

Hansen, Donald B.
1983 *Generic GIS (Geographic Evaluation Network for Earth Resource Inventory and Classification).* Western Energy Land Use Team, Division of Biological Services. U.S. Fish and Wildlife Service, Ft. Collins, Colorado.

Hart, Judy A., and David B. Wherry, editors
1984 *VICAR/IBIS User Reference Manual.* Washington State University Computing Service Center, Pullman.

Hasenstab, Robert J.
1983a The Application of Geographic Information Sysrems to the Analysis of Archaeological Site Distributions. Paper presented at the 48th Annual Meeting of the Society for American Archaeology, Pittsburgh.

1983b *A Preliminary Cultural Resource Sensitivity Analysis for the Proposed Flood Control Facilities Construction in the Passaic River Basin of New Jersey.* Soil Systems, Inc. Submitted to U.S. Army Corps of Engineers, New York District. Copies available from Soil Systems, Inc., Marietta, Georgia.

Hodder, Ian, and Clive Orton
1976 *Spatial Analysis in Archaeology.* Cambridge University Press, Cambridge, England.

Hoffer, R., and staff of LARS
1975 *Computer-Aided Analysis of Skylab Multispectral Scanner Data in Mountainous Terrain for Land Use, Forestry, Water Resources, and Geologic Applications.* LARS Information Note 121275. Laboratory for Applications of Remote Sensing (LARS), Purdue University, Lafayette, Indiana.

Jochim, Michael A.
1976 *Hunter-Gatherer Subsistence and Settlement: A Predictive Model.* Academic Press, New York.

Johnson, G. A.
1977 Aspects of Regional Analysis in Archaeology. *Annual Reviews in Anthropology* 6:479–508.

Kvamme, Kenneth L.
1983a *New Methods for Investigating the Environmental Basis of Prehistoric Site Locations.* Ph.D. dissertation, University of California. University Microfilms, Ann Arbor.

1983b Computer Processing Techniques for Regional Modeling of Archaeological Site Locations. *Advances in Computer Archaeology* 1:26–52.

1984 Models of Prehistoric Site Location Near Pinyon Canyon, Colorado. In *Papers of the Philmont Conference on the Archaeology of Northeastern New Mexico,* edited by Carol J. Condie, pp. 349–370. New Mexico Archaeological Council Proceedings 6(1). Albuquerque.

1985 Geographic Information Systems Techniques for Regional Archaeological Research. Paper presented at the International Symposium on Data Management and Mathematical Models in Archaeology, International Union of Pre- and Protohistoric Sciences, Denver.

1986 The Use of Geographic Information Systems for Modeling Archaeological Site Distributions. In *Geographic Information Systems in Government,* vol. 1, edited by Bruce K. Optiz, pp. 345–362. A. Deepak, Hampton, Virginia.

Leberl, Franz W., and Dale Olson
1982 Raster Scanning for Operational Digitizing of Geographical Data. *Photogrammetric Engineering and Remote Sensing* 48:615–627.

Lee, John E., Carl N. Reed, Randy L. Frosh, and LeaAnn Sullivan
1984 *MOSS Users Manual, Version 84.07.* Western Energy and Land Use Team, Division of Biological Services. U.S. Fish and Wildlife Service, Washington, D.C.

Lewis, Peter
1977 *Maps and Statistics.* Methuen, Cambridge, England.

Loveland, Thomas R., and Gary E. Johnson
1983 The Role of Remotely Sensed and Other Spatial Data for Predictive Modeling: The Umatilla, Oregon Example. *Photogrammetric Engineering and Remote Sensing* 49:1183–1192.

Lyons, Thomas R., and Robert K. Hitchcock
1977 *Aerial Remote Sensing Techniques in Archeology.* Reports of the Chaco Center No. 2. National Park Service and University of New Mexico, Albuquerque.

MacDougall, E. Bruce
1971 *Computer Programming for Spatial Problems.* John Wiley, New York.

Marble, Duane F., Hugh W. Calkins, and Donna J. Peuquet
1984 *Basic Readings in Geographic Information Systems.* SPAD Systems, Williamsville, New York.

McLeod, Ronald G., and Beverly K. Jafek
1984 Solving Environmental Problems with Satellite Imagery: Use of the Geographic Information System at JPL. *Computer Graphics World* 7:75–78.

Missallati, A., A. E. Prelat, and R. J. P. Lyon
1979 Simultaneous Use of Geological, Geophysical, and Landsat Digital Data in Uranium Exploration. *Remote Sensing of Environment* 8:189–210.

Moellering, Harold
1980 The Real-Time Animation of Three-Dimensional Maps. *The American Cartographer* 7:67–75.

Moik, Johannes G.
1980 *Digital Processing of Remotely Sensed Images.* National Aeronautics and Space Administration, Washington, D.C.

Monmonier, Mark S.
1974 Measures of Pattern Complexity for Choropleth Maps. *American Cartographer* 1:159–169.

1982 *Computer Assisted Cartography: Principles and Prospects.* Prentice-Hall, Englewood Cliffs, New Jersey.

Munro, Patricia A.
1983 Selection Criteria for a Geographic Information System—In Retrospect. In *Renewable Resource Inventories for Monitoring Changes and Trends,* edited by John F. Bell and Toby Atterby, pp. 273–275. Oregon State University, Corvallis.

National Cartographic Information Center
1980 *Digital Terrain Tapes: User Guide,* 2nd ed. U.S. Geological Survey, Washington, D.C.

National Research Council
1983 *Procedures and Standards for a Multipurpose Cadastre.* National Academy Press, Washington, D.C.

Parker, Sandra
1985 Predictive Modeling of Site Settlement Systems Using Multivariate Logistics. In *For Concordance in Archaeological Analysis: Bridging Data Structure, Quantitative Technique, and Theory,* edited by Christopher Carr, pp. 173–207. Westport, Kansas City.

1986 The Role of Geographic Information Systems in Cultural Resource Management. In *Geographic Information Systems in Government,* vol. 1, edited by Bruce K. Optiz, pp. 133–140. A. Deepak, Hampton, Virginia.

Reed, Carl N., III
1986 DELTAMAP—Just Another New GIS? In *Proceedings, Second International Symposium on Spatial Data Handling, July 5-10, 1986, Seattle, Washington* pp. 375-383. International Geographical Union, Commission on Geographical Data Sensing and Processing, Williamsville, New York.

Rhind, David
1975 A Skeletal Overview of Spatial Interpolation Techniques. *Computer Applications* 2:293–309.

1981 Geographic Information Systems in Britain. In *Quantitative Geography: A British View,* edited by N. Wrigley and R. J. Bennett, pp. 17–35. Routledge and Kegan Paul, London.

Rogers, Alan, and John A. Dawson
1979 Which Digitizer? *Area* 11:69–73.

Scholtz, Sandra Clements
1981 Location Choice Models in Sparta. In *Settlement Predictions in Sparta,* by Robert H. Lafferty III, Jeffrey L. Otinger, Sandra C. Scholtz, W. Frederick Limp, Beverly Watkins, and Robert D. Jones, pp. 207–222. Arkansas Archaeological Survey Research Series No. 14. Fayetteville.

Schowengerdt, Robert A.
1983 *Techniques for Image Processing and Classification in Remote Sensing.* Academic Press, New York.

Shelton, Ronald L., and John E. Estes
1981 Remote Sensing and Geographic Information Systems: An Unrealized Potential. *Geo-Processing* 1:395–420.

Steiner, D., and A. E. Salerno
1975 Remote Sensor Data Systems. In *Manual of Remote Sensing*, edited by Robert G. Reeves, pp. 611–803. American Society of Photogrammetry, Falls Church, Virginia.

Stow, Douglas A., and John E. Estes
1981 Landsat and Digital Terrain Data for County-Level Resource Management. *Photogrammetric Engineering and Remote Sensing* 47:215–222.

Strahler, Alan H., Thomas L. Logan, and Nevin A. Bryant
1978 Improving Forest Cover Classification Accuracy from Landsat by Incorporating Topography Information. *Proceedings of the Twelfth International Symposium on Remote Sensing of Environment*, pp. 927–974. Environmental Research Institute of Michigan, Ann Arbor.

Tesche, T. W., and R. W. Bergstrom
1978 Use of Digital Terrain Data in Meteorological and Air Quality Modeling. *Photogrammetric Engineering and Remote Sensing* 44:1549–1559.

Thompson, R. A.
1978 *A Stratified Random Sample of the Cultural Resources in the Canyonlands Section of the Moab District of the Bureau of Land Management in Eastern Utah.* Cultural Resource Series No. 1. Bureau of Land Management, Salt Lake City, Utah.

Tomlinson, R. F., H. W. Calkins, and D. F. Marble
1976 *Computer Handling of Geographical Data: An Examination of Selected Geographic Information Systems.* Natural Resources Research XIII. UNESCO, Paris.

Turner, A. Keith
1978 A Decade of Experience in Computer Aided Route Selection. *Photogrammetric Engineering and Remote Sensing* 44:1561–1576.

Wehde, Mike
1982 Grid Cell Size in Relation to Errors in Maps and Inventories Produced by Computerized Map Processing. *Photogrammetric Engineering and Remote Sensing* 48:1289–1298.

Woodcock, Curtis E., Alan H. Strahler, and Thomas L. Logan
1980 Stratification of Forest Vegetation for Timber Inventory Using Landsat and Collateral Data. In *Proceedings of the Fourteenth International Symposium on Remote Sensing of Environment*, pp. 1769–1787. Environmental Research Institute of Michigan, Ann Arbor.

Yoeli, P.
1975 Compilation of Data for Computer-Assisted Relief Cartography. In *Display and Analysis of Spatial Data*, edited by John C. Davis and Michael J. McCullagh, pp. 352–367. John Wiley, New York.

Zobrist, A. L., and N. A. Bryant
1979 *Image-Based Information Systems (IBIS)—System Guide.* Technical Report 900-909. Jet Propulsion Laboratory, Pasadena.

Zubrow, Ezra B. W.
1971 Carrying Capacity and Dynamic Equilibrium in the Prehistoric Southwest. *American Antiquity* 36:127–138.

Chapter 11

PREDICTIVE MODELING AND ITS RELATIONSHIP TO CULTURAL RESOURCE MANAGEMENT APPLICATIONS

Chris Kincaid

One goal of this volume, stated in Chapter 1, is to "explore the feasibility and practicality of predictive modeling for meeting management objectives." We will address this goal in the following pages. First, however, we need to consider just what our management objectives are, and how they relate to what might be called "research objectives."

Research objectives of modeling tend to fall under the general heading of "an improved understanding of the archaeological record." Models can improve our definition and recognition of important types of sites and our understanding of their distribution across the landscape. Models can clarify processes of culture change and interaction and provide a regional framework for understanding the development and evolution of human systems. They can permit us to understand cultural adaptation to differing environments and provide insight into the nature and origin of social, political, and economic processes.

While initially such information might seem abstract and removed from the practical requirements of cultural resource management, in reality it meets several critical management objectives. Management objectives are sometimes thought to be limited to a narrow concern over "how many sites are where," and indeed, models can suggest what types of sites are in a specific area and where in that area they might occur. Some models can also be used to generate population estimates and statements concerning the probability of site occurrence in a particular location. These classes of information are important in management decisions about possible surface-disturbing actions. But the more research-oriented objectives of modeling are also important because such models can help to indicate data gaps and highlight research issues needing additional work. In this way the use of models can help us to focus scarce agency dollars on the collection of the most necessary and *important data* and reduce waste caused by repetition. Such models can help us to learn more from existing data and, in some cases, can expedite and streamline inventory programs. While some products or applications of models are more important in either a research or a management context, in a broader sense research and management objectives overlap a great deal, and both stand to gain from a model that is reliable and adequately explains as well as predicts site occurrences.

In the following pages we will explore the application of specific theoretical and methodological approaches described in preceding chapters to modeling in a management setting. The discussion is organized around the topics of preparing for, implementing, and evaluating a modeling project. Practical considerations are foremost. The goal here is to highlight the benefits of using modeling in cultural resource management, while at the same time indicating some of the potential limitations of its use. The desired end result is a balanced and responsible application of modeling concepts to management situations. Although discussions are keyed to land management issues, we have tried not to limit them to a single-agency perspective.

Questions have been raised as to whether inventory and evaluation strategies employing modeling techniques meet the intent of the National Historic Preservation Act, Section 106. This legislation requires a determination of the effect of federal agency actions, or federally permitted or licensed actions, on all properties listed in or eligible for listing in the National Register of Historic Places.

Under the provisions of this legislation, decisions about appropriate inventory and evaluation strategies are made through consultation between the federal agency and the State Historic Preservation Officer (sometimes also including the Advisory Council on Historic Preservation). There are no set criteria for deciding what is appropriate; rather, propriety is defined on a case-by-case basis through the consultation process, within the broad structure of the regulations. The decision as to whether or not modeling should be part of an inventory and evaluation approach depends on individual circumstances. A decision to use modeling complies with the regulations if it was reached in accordance with the consultation procedures. For this reason, compliance questions are not addressed further in this chapter.

WHAT ARE MODELS ABOUT?

As analytical tools, archaeological resource models are especially well-suited to applications in land management. Among other things, they identify patterns in spatial relationships between sites and their physical locations and thus indicate potential relationships between the natural or social environment and the locations of past human activities. A causal relationship is envisioned: environmental factors influence where human activities occur. Measurements that define or describe controlling aspects of the natural or social environment are called independent variables, while measurements of affected human activities, observable in the archaeological record, are called dependent variables.

The development of models centers around three main tasks: classification of independent variables, classification of dependent variables, and expression of the relationship between them. Since different cultural groups interact with each other and their environment in different ways, the critical independent and dependent variables and their relationship can vary widely from cultural system to cultural system. The goal of this kind of modeling is to produce reasonably accurate

representations of selected interrelationships for particular cultural systems. A successful model or series of models allows us to organize what we know about sites—their function, location, and cultural affiliation—into a series of affirmative statements about human behavior. Under controlled conditions, these statements can be applied to unknown areas to provide predictions about resources located in these areas.

Our goal is to correctly identify important aspects of the natural or social environment that influenced the location of human activities, and to interpret the archaeological record as the result of a set of functional, temporal, spatial, and behavioral responses to a varied environment. We may, in effect, try to reconstruct the "rules" of interaction between these two components. The relationship between sites and their natural environments is not as easily discoverable or as direct as the relationships among natutal phenomena. Although governed to an extent by the demonstrably regular and consistent rules that apply to all living systems, human behavior is organized into cultural systems, which exert additional influences on that behavior beyond those of natural forces. There is good reason to believe that site locations cannot, in general, be fully predicted from environmental variables alone.

Because of the influence of cultural variables on human behavior, models of cultural systems are subject to many more sources of error than those for natural systems. The cultural rules that govern how human groups interact with their social and natural environments are not easy to identify, even for modern cultures. Studying and identifying such relationships for cultures that have been extinct for thousands of years is an even more difficult task. In land management applications, therefore, models of natural phenomena and models of cultural phenomena should not be considered equivalent. Managers need to have a realistic understanding of what models can and cannot do in order to use them effectively.

WHAT CONDITIONS ARE FAVORABLE FOR MODELING PROJECTS?

Conditions

Before a decision is made to embark on a modeling project to satisfy either research or management objectives, several conditions must be met. Frequently, these conditions relate to circumstances, such as the boundaries of the study area, time, financial constraints, etc., that are beyond the control of the project manager. For example, the size of the potential modeling project area is important. As a general rule, modeling is not feasible for small projects covering less than 5000 to 10,000 acres. Models are most easily interpreted and understood if they relate in a defined way ro cultural boundaries or to major environmental zones. When only a small portion of a culture area or environmental zone can be analyzed, only a portion

of a cultural system might be examined. Observed site patterning in the study area may be responding to factors that are "uncontrolled" in the terms of the model because they are a response to forces or events located outside the study area. The chances of developing an accurate and interpretable model are greatly reduced by this circumstance.

In designing small modeling projects, difficulties often occur in meeting minimum sample-size requirements for statistical analyses. Altschul and Nagle address this problem in some detail in Chapter 6. In general, they advise that for cluster analysis of site types, a minimum of 30 sample units (not including empty units) is necessary, a condition frequently not met in sample inventories. Unfortunately, the number of sample units necessary for a valid analysis cannot be anticipated prior to fieldwork. Additional inventory may be required to reduce sample variance if a majority of sample units do not contain sites while the remaining units contain many sites.

The configuration of the modeling project area is also important. Linear as compared to areal projects are generally more difficult to model because linear projects tend to cross-cut several environmental and cultural zones, each of which may be poorly represented as regards total acreage. More complex models, or additional models, may be needed in these cases.

Another important factor is the amount of time allotted for the modeling project. Modeling is useful as a long-term technique for organizing and structuring data and data collection priorities. It is less useful under a short time frame that does not allow for testing and refinement phases.

Often the nature of the archaeological record itself can indicate that special strategies will be needed for modeling efforts. For example, if 75 percent of the known sites in an area are classified as undiagnostic lithic scatters, neither chronologically nor functionally specific models can be developed. Under these circumstances, care should be taken in designing any new sample inventories in the area to assure that detailed information pertaining to attributes of artifacts is collected. This data could be crucial in the definition of site classes during postinventory modeling efforts.

Sometimes the environment determines whether modeling will be easy or difficult. In Chapter 4, Ebert and Kohler distinguish between environmental variables (which measure a single aspect of the environment, e.g., slope) and ecosystemic variables (which measure systemic attributes reflective of interaction among environmental variables, e.g., effective temperature, spatial periodicity, and environmental diversity). The most usable ecosystemic variables for predicting site locations are those that monitor spatial availability of resources (e.g., degree of patchiness) and temporal availability of resources (e.g., degree of constancy, contingency, and predictability). Ebert and Kohler conclude that, in general, heterogeneous environments in which critical resources are temporally predictable and occur in highly concentrated and overlapping patches are apt to be best for locational modeling and prediction. Conversely, a basically homogeneous environ-

ment, in which critical resources are dispersed and only sporadically available, will be more difficult to model. For example, site locations may be more difficult to model on a desert creosote flat without major drainages or contrasting landforms than they would be on a flat broken by large dry washes or benchlands, where critical vegetative resources are apt to occur.

Also, changes in the earth's surface may have taken place after the deposition of archaeological materials (see discussion of postdepositional processes in Chapters 4, 6, and 9). Postdepositional processes might include movement of sand dunes, deposition of alluvium, erosion by wind or water, etc. Large portions of a project area may have been covered over or scoured away as a result of these processes. In areas undergoing active alluviation, for example, exposed surfaces may be no older than 200-300 years. The possibility of finding prehistoric sites under such conditions is greatly reduced, and modeling efforts directed to prehistoric site locations would be unproductive. Under these circumstances specialized strategies (such as inventories focused on road cuts, arroyos, etc.) may be appropriate.

Administrative Concerns

The risks of embarking on a modeling project should be evaluated realistically at the onset of a project and weighed against such administrative constraints as project schedules and costs. To develop a model that meets a specified level of precision, additional testing and analyses may be needed, sometimes causing delays and increased costs. Clearly, the importance of these concerns will depend on the type of modeling project envisioned and its use.

Time should be allowed for model testing and revision during any project. In Chapter 6, Altschul describes a multistage survey design, a useful means for staging sample-based fieldwork so that the maximum benefit is derived from each successive stage. While a multistage approach may seem more time-consuming, experience has shown that a single data-collection phase is seldom adequate for model development (depending, of course, on the size of the initial sample and the availability of relevant historical, ethnographic, and other data) and may be less efficient over the long term.

If a model is being developed to reduce the cost of field inventory, various hidden costs should be taken into account. Short-term field inventory costs are almost always less for partial coverage than for full coverage, even allowing for the substantial field time needed to locate dispersed sample units, but the cost of developing a predictive model is not limited to the costs of the sample inventory. In any given modeling project, time and funds also should be allocated for such tasks as

1. the detailed analysis of existing information,
2. preparation of environmental data,
3. development and execution of successive phases of model testing (using independent data), and

4. collection and processing of supplemental information about site variability (through various combinations of detailed recording, surface collection, testing, excavation, and laboratory analyses).

Planning for these additional costs is not easy. Exact estimates as to the amount of work that will be needed to develop a model of the required precision cannot be made before a project is even begun.

Perhaps the most cost-effective context for model development is within the framework of general planning by a land-managing agency or a local government. These programs can develop and sustain long-term approaches that are funded incrementally and result in cumulative and refined data bases. Such data bases, and the models based on them, may take years to develop and test. The end result, however, is a powerful and effective management tool.

WHAT KINDS OF MODELS ARE THERE? WHEN DO WE USE WHICH TYPE?

Models are classified in many different ways. In Chapter 2, for example, models are compared with respect to their focus (systemic, representing a cultural system; analytic, reflecting the analysis of archaeological data), their logical origin (inductively or deductively derived), and the level of measurement (nominal, ordinal, interval, and ratio scales). Figure 2.1 presents the structure of this discussion.

In Chapter 3, intuitive models are distinguished from objective models on the basis of whether or not components can be operationalized or measured. Objective models are then broken down further on the basis of geographic precision (are predictions specific to points or to areas?), on procedural logic (inductive versus deductive reasoning), and on the relative emphasis given different variables. A summary of this approach is presented in Table 3.1.

Chapter 5 contains a discussion of various statistical techniques that have been used to classify models (e.g., linear regression, logistical regression, and discriminant function analysis). Kvamme, in Chapter 8, distinguishes between models based on trends in "location only" (defined solely in terms of locational coordinates x and y) and models based on trends in "locational characteristics," or a wide range of environmental attributes of these locations. He further divides models pertaining to the characteristics of locations according to whether they are based on parametric or nonparametric statistics.

How does the cultural resource manager know which type of model is best? Is it not possible to define one type of model that is best for cultural resource management purposes and apply this type to all situations?

To understand the significance of the modeling terms used in various portions of this volume, we should view them not as designations of types of models but as descriptive labels for various traits or attributes of models. A cultural resource

manager seeking the best model for his or her purposes must ask first, *what is the overall objective of developing a model?* If there is a need that only one type of model will fill, then clearly this type of model must be sought. More often, a manager simply seeks the most precise, detailed, extensible, and accurate model affordable. Such considerations as the nature of the existing data base, environmental complexity, etc. (see discussion in the previous section), ultimately intervene to limit the quality of the model obtainable with given data.

One of the broadest and most useful classifications is the one based on procedural logic, distinguishing between inductive and deductive approaches. The relative merits of these strategies have been debated throughout this volume. Briefly, a deductive approach, i.e., one proceeding from theory to data, often explains why a model works. This is necessary, especially if the model is to be successfully applied to other settings. The major drawback of deductive models is the difficulty in making them operational. For example, deductive models often contain general propositions, such as "population growth leads to more intensive resource utilization." The archaeologist must determine how "population" and "resource utilization" will be measured to show growth and increased intensity. Abstract concepts such as these may be difficult to measure in tangible terms from archaeological data, especially if these data are sparse, as is often the case.

In contrast, inductive models proceed from data to theory; observed correlations in the data are used to formulate general hypotheses. If, for example, several major village sites in a particular area are located near or on one particular soil type, one might hypothesize that large habitation sites tend to be located close to this particular soil type. Such conclusions may be readily derived through data analysis, but models that depend on them are often criticized for not explaining why the observed correlations occur. Most models developed for cultural resource management purposes are inductive.

Clearly, managers should understand why a model works, but in addition they need an approach that is operational. Joseph Tainter (personal communication, 1982), in commenting on one of the initial drafts of this volume, offered the following observations on this matter:

> The crucial question is not whether a model is derived deductively or inductively, but whether it focuses on explaining patterns or merely projecting them. Explanations can precede or follow data collection, but must be developed at some point.

One way of achieving this may be to structure the modeling process to be sure that both deductive and inductive phases are included.

In reality, in the long-term time frame of cultural resource management programs, the distinction between deductive and inductive approaches becomes blurred. The model building and refinement process is based on a continuous cycle of data collection, analysis, and model refinement. The results of one cycle of field testing and analysis are used to refine the model, which then guides the next phase of data collection. The eventual merging of deductive and inductive strategies may

be the direction of future modeling approaches in cultural resource management contexts.

How does a manager know which specific type of model is needed for a particular application? If a model is needed for a limited, short-term application to a relatively small project area (for example, in connection with the processing of a right-of-way or an energy development application), a relatively limited range of models will be appropriate. For more general, long-term applications in the cultural resource program, a wider range of models could be applicable and useful.

Given the complexity of cultural resource management, virtually every type of model has some utility. Both inductive and deductive models are appropriate in varying degrees, depending on the circumstances. Deductive models have the greater utility in developing inventories and in such program activities as site interpretation. Inductive, correlative models usually have the statistical precision needed to develop quantitative estimates of site populations, densities, and distributions and are currently the better source of such estimations. Both types of models may be needed in a comprehensive cultural resource management program.

HOW CAN WE PREPARE FOR MODEL DEVELOPMENT?

Model development is a repetitive process of inventory and analysis that is most effective as a long-term strategy. In general, the quality of the model depends on the quality of the data; better data bases yield more precise and accurate models.

Even before beginning the modeling process, the cultural resource specialist can take many steps that do not require large-scale or expensive sample inventories. Since the beginning of cultural resource management programs, managers have recognized the need to make full use of existing information. Chapter 7 specifically addresses model-building requirements and techniques to develop good data bases. As a first step, the cultural resource specialist should accumulate and screen all available information on the study area's history and ethnography, and on previous survey work in the area. The quality of data on previously recorded archaeological sites and other historical properties should be carefully reviewed for locational accuracy and completeness, and sites for which information is poor should be set aside for later evaluation. Checking selected sites in the field may be necessary to evaluate recording practices and improve information.

The second step should involve assembling information into a coherent, usable format. If the equipment and expertise are available, this might include automating the site data base. Several regions and states have systems for managing site data. If one of these is not available, a data base can be set up on an office computer. Automating the data base allows the specialist to review the data easily and informally, and to evaluate them apart from any ongoing modeling project. Analyses of existing data for future modeling projects will be much simpler and less expensive than current methods. Subfiles can be used to store more detailed site

and artifact data specific to individual sites. These subfiles can be created easily and accessed as needed during detailed analysis.

All surveyed areas should be mapped on base maps. The type and completeness of survey coverage must be carefully scrutinized. Using information provided in project reports, the specialist should separate projects in which coverage appears to have been biased, incomplete, or otherwise suspect from those in which survey and recording practices conform to acceptable standards. Only survey data that are relatively complete can be used confidently in studies of spatial distribution. While data recorded during less rigorous, nonstandard surveys may be very useful for site-level analyses, methodological biases that distort apparent spatial distributions make these data unsuitable for modeling purposes.

Documents summarizing existing data in an area, such as Class I inventories, state plans, and regional research designs, can be especially useful as a preliminary source of site distribution information within a study area. Research designs should summarize what is known about an area in the form of model-like statements or hypotheses, which can then be tested when new data are collected. These studies can be completed on a contract basis, generally at relatively low cost.

The definition of site types reflecting temporal, functional, and cultural differences is perhaps one of the most useful tasks that can be performed to prepare for model building. (Procedures for this task are discussed in Chapters 5 and 8.) Site types or other similar classification schemes are one of the primary components of models.

Environmental data are also needed for model building. To be useful, however, they should be of a consistent quality and scale throughout the study area. Land-managing agencies typically expend considerable effort in collecting a wide range of environmental data for land-use planning and environmental impact considerations. This is done through field inventories, analysis of aerial photographs and other remote sensing data, GIS development, etc. The manager should ensure that such data collection projects take into consideration the unique needs of the cultural resource program. These needs (e.g., for data pertaining to the paleoenvironment or identifying postdepositional processes) should be anticipated by the manager and, where possible, collected as part of other specialized studies. In areas of adjacent or mixed jurisdiction, opportunities for interagency development of environmental data can be explored to reduce costs.

Once the requisite data bases have been assembled, screened, and organized, several kinds of preliminary analyses can be performed to evaluate and characterize the data. This step is actually the beginning of the model-building process, which will be discussed further in the following section. These preliminary analyses are apt to be biased and inaccurate, however, because the existing data used at this point probably do not represent the study area as a whole.

This does not mean that trial models developed at this stage are unusable, only that their use is limited, and that they should be used with caution. Trial models provide a check on the adequacy of field recording procedures. Even an initial

modeling exercise may point out the need for additional detail in artifact recording, for subsurface testing, for an increased level of site examinations, for a shift to interval or ratio-scaled data, etc.

Appropriate changes made in site recording techniques can greatly increase the chances of successful modeling efforts in the future. Even deficient models can help to identify gaps in inventory coverage and highlight new data needs. Similarly, future large-scale modeling efforts may improve substantially if preliminary small-scale projects can first be applied to areas for which we have relatively little information.

Statistically representative data are not necessary to develop a model; if new data collection *is* planned for purposes of model development, however, these are certainly the most effective data to collect. Model *testing*, on the other hand, does depend upon the availability of unbiased data that are representative of the study area, most often data that were collected using some form of random sampling. Until a representative sample of data is obtained through a carefully designed inventory project, any model developed for the area must remain essentially untested and should be used accordingly (see the section on model evaluation, below).

HOW DO WE PLAN A MODEL?

Should funding become available for a modeling project, several measures can be taken to ensure that management needs will be met and that the project will be as successful as possible. To begin with, in what might be called a preplanning phase, the goal needs to be clearly defined. The purpose of the project should be carefully considered, recorded, and reviewed by managers and other resource-program staff members. Both long-term and short-term goals should be considered, including all potential management applications of the product, as well as immediate uses. The possibility of phasing the project over a period of years should be considered, depending on whether one-time or continued funding is anticipated.

Two important decisions to be made are the size of the target study area and the type and resolution of desired model products. For large study areas, entering into joint projects with agencies or others (e.g., Indian tribes or local governments) who manage adjacent lands may be advantageous, especially if the combined land base more nearly addresses a meaningful cultural or environmental unit. Establishing two study areas may also be advantageous—a larger one to be used during the analysis of existing data and a smaller, more limited one to be used in definition of the target population of inventory sample units.

The full range of available model products and their limitations should be weighed to ensure that initial expectations match the results. Possible types of products include statements about relative site distributions, population estimates (e.g., estimated numbers of sites in unsurveyed sample units, numbers of unsur-

veyed sample units without sites, and total site population), correlations between certain environmental factors and the locations of certain types of sites, and probability statements (e.g., the probability of finding a site in any one sample unit or the probability that the observed result would occur by chance alone).

In many cases, the value of the results depends on the detail of the environmental data recorded for each sample unit and the levels and types of measurement used in recording the data. These factors should be addressed during the planning phase of the modeling projecr and should reflect the management constraints, goals, and limitations identified during the preplanning phase.

After the management needs have been clearly defined, a modeling project plan should be developed. The purpose of a project plan is to break down the modeling process into a series of review and decision points, whereby the manager or specialist and the individual implementing the project (in most cases a contractor) can review progress and jointly participate in key decisions.

The first step in the development of the plan should be a review of existing data and formulation of a trial model for the study area. Plans to test this trial model should be described in a research design that clearly spells out research issues, data gaps, and priorities for collection of new data during sample inventory work. Statements should address data-collection activities: selection of inventory areas; data recording for micro-and macro-environmental data; and recording of site, feature, and artifact attributes. Each recording activity should be carefully reviewed to ensure that the most powerful measurement system will be used, a critical factor if inventory results are to have the maximum applicability to modeling efforts.

Detail concerning the rationale for selection of sample inventory units should be provided in the sample design. The sample design should clearly describe any proposed stratification schemes and their goals, plus the configuration of the sample units and the method of their selection. It should also evaluate the need for multiple survey strategies (e.g., a mix of random and judgmental samples).

At a minimum, the results of the initial data review, the trial model, and the data-collection proposal should be included in a preliminary report prepared prior to initiation of fieldwork. The manager and specialist can thus determine at this preliminary stage whether maximum use has been made of existing data and can ensure that the first stage of field inventory is directed toward a model testing effort. Peer review of this report may be desirable.

The plan should next address the second step of the project—the fieldwork phase. Detailed information about the proposed field methods, including rates of inventory, recording standards, collection strategies, and schedules, should be provided.

The third step of the project plan should address analysis and preparation of the final report. In this part of the plan, proposed approaches to data preparation and analysis should be described. The relationship between proposed products of

analysis and the original management goals should be discussed, even though at this stage the former might be tentative. Ultimately, the results of fieldwork—including the number, variability, and distribution of sites—will have a principal role in determining the level of analysis possible. The anticipated artifact studies or other laboratory analyses designed to distinguish site types and functions should be described as well.

The final report should be structured to present the model, describe the uses of the data, explain differences between the initial trial model and the final model, and describe its limitations and appropriate applications. An explicit statement should also be included to detail how the model could or should be tested, regardless of whether additional inventory is envisioned in the near future.

Specific technical information on the most effective ways of performing modeling tasks has been presented throughout this volume; this information should be read carefully by any cultural resource specialist responsible for overseeing or monitoring a modeling project. A description of the overall modeling process from the perspective of the land managing agency is provided by Altschul in Chapter 3. A critical discussion explaining types of measurements and their importance to modeling is presented in Chapter 5, followed by an extensive treatment of the mechanics of the model-building process, including development of site classifications. Specific topics such as sampling strategies, parameter estimation, the empty unit problem, phased sampling and survey, and data-recording strategies are covered in Chapter 6 by Altschul and Nagle. Kvamme, in Chapter 7, analyzes the use of existing data in trial model formulation. In Chapter 8 he looks more closely at different types of models and compares their output and applicability to management situations. Chapter 8 also contains a review of techniques for model testing and refinement, addressing such topics as parametric and nonparametric statistical analyses and assumptions about the data, testing, and confidence intervals.

HOW DO WE APPLY MODELING IN CULTURAL RESOURCE MANAGEMENT?

Many aspects of cultural resource management can benefit directly or indirectly from the use of modeling techniques. Even if a formalized model is not developed, the techniques used to prepare cultural resource data for a modeling exercise (see the section on preparation for model development, above) can have useful side benefits. Some of these are discussed below.

Inventory

Models can be used in the design of comprehensive inventories specific to a defined land base or land-use area. Within this area, information concerning site

types, site locations, and environmental characteristics can be ordered as either dependent or independent variables and (under an inductive approach) analyzed to detect patterning in the data. Through this process, additional data needs (e.g., for more detailed site data, for inventory coverage in specific places) become evident and can be prioritized. Even in preliminary modeling efforts (deductive or inductive), new types of information, such as siteless locations, paleoenvironmental data, and information on postdepositional processes, are often needed.

Models can also drive inventory efforts when information is needed to *test* (rather than develop) the model. While models can be created from a diverse and not rigorously representative data base, they can only be tested properly using an unbiased data base, i.e., one that represents the study area as a whole. Because of the overall utility of models, the collection of sample inventory data for model testing and refinement should be a high inventory priority, regardless of whether these data are selected within an administrative land base (e.g., a resource area or forest) or a limited study area within that administrative land base.

Because of limited funding, some managers may consider turning to modeling as a substitute for, or as a way of limiting, new inventory data collection. This is not a cure-all approach, however, because the results of a model are only as good as the data on which the model is based. For this reason, models sometimes do a poor job of predicting variability and may not be reliable or precise. While each case must be evaluated on its own merits, there are several criteria that cultural resource managers should consider when deciding how to use modeling in field inventory efforts. One important consideration is the possible repercussions if scarce inventory dollars are spent to develop a model that cannot perform to the desired level of accuracy. The purpose for which an inventory is conducted will determine how serious this problem will be. The type of model being used must be evaluated with respect to the application under consideration. The analytical origin of the model is important, as is the question of whether it has been tested. (More specific criteria for model evaluation are included in the next section.)

The types of sites in the study area are also an important factor. It is one thing to limit inventory in an area thought to contain homogeneous archaeological remains, such as small sites with limited variability, no depth, and shared attributes. It is quite different to limit inventories in an area known to contain complex, large, or stratified sites; a heterogeneous site population; or what Altschul and Nagle (Chapter 6) refer to as *magnet* sites (sites thought to influence the location of other sites).

The scale of resolution of the model is important. Zonal models perform differently from point models and generally cannot provide specific site-likelihood indications for designated locations.

Land managers should guard against the improper use of intuitively derived models in influencing inventory efforts. Archaeologists who work frequently in an area often develop a "feel" for where sites should be found. Occasionally, these intuitions have been used as a basis for limiting inventory to certain areas without

testing others. A danger in this approach is that if sites are sought only where they are thought to exist, the prediction may become a self-fulfilling prophecy. Potential results can include destruction of significant resources or introduction of a strong bias into the data base.

Intuitions should not be dismissed, but neither should they be equated with scientifically verified information. They should be formalized, expressed in terms that can be measured and applied in the inventory process, and subjected to a rigorous testing program. In this way they can be of vital importance in effective model development.

Evaluation

In evaluating an archaeological site for management purposes, two major kinds of questions are asked. One has to do with a site's significance (generally expressed in terms of eligibility for the National Register of Historic Places). The other has to do with determination of its most appropriate use(s), toward which further management actions should be directed. Modeling can contribute to both kinds of evaluation.

The significance of a site can be measured by its potential to contribute to our understanding of the historical and prehistoric past. On a broad scale, models can help to clarify these research issues, thus providing a more consistent regional context for site evaluation.

By focusing research on the location of sites, as well as on the *types* of sites expected to occur in specific locations, the modeling process can help to increase the accuracy and precision of functional, temporal, and spatial qualifiers. Modeling helps to define major similarities and differences among sites and reflects the information potential for both identified and projected sites within an area. In evaluating whether a particular site is potentially significant, the specialist often relies on previous experience with other sites of the same type.

The importance of a site cannot be equated solely with its membership in a particular site type or class, however; clearly, the rare or unique site, which fails to appear as a separate type during statistical analyses, may be the most significant site in an area. These sites are often not amenable to identification through sample inventory, but they can be successfully integrated into predictive models if sufficient information is known about them. This issue is discussed further in the next section.

It is important to consider the physical characteristics of a site as well as its class membership. For example, a broad class labeled "habitation sites" might include sites with or without structures. Some large sites might be in poor condition, with virtually no remaining information potential, while some small sites might contain substantial undisturbed deposits. Clearly, significance assessment must address individual site characteristics, as well as class membership.

Detailed information about the relative scarcity, relative research importance, and locations of archaeological sites, whether they have all been discovered and recorded or not, can help in the determination of their appropriate uses. Examples of possible uses include ongoing or potential scientific study, maintenance of a social and/or cultural group's traditional lifeways, public education and interpretation, and experimental management studies. By refining ideas about all of the archaeological remains in a target area, models can be extremely valuable for focusing and organizing use and allocation decisions.

Protection

Another important program component where modeling can be applied is in the area of protection. Protection refers to measures taken to reduce natural or human-caused impacts to the significant qualities of cultural properties or to the attainment of their appropriate uses. Measures may include information signs, physical barriers, patrol and surveillance, monitoring, detailed recording, excavation, stabilization, and administrative measures, such as access restrictions, withdrawals from other land-use activities, avoidance during construction, etc.

The principal way in which modeling can contribute to protection is by helping to establish priorities among sites for specialized treatment. In general, land-managing agencies are interested in protecting and preserving a "representative array" of sites and site data. It is useful to visualize this array in terms of types of sites; all sites of a like type constitute a finite site pool. Theoretically, protection and preservation efforts should be directed toward maintaining a representative site pool of each site type for future needs. Modeling provides a basis for determining the array of site types in a particular area, as we have seen, and in some cases can be used to generate population estimates for various site pools.

Models can also help to define research issues. This information can serve to guide data collection priorities for data recovery efforts and can help to establish which sites should be selected for these efforts. Models can be used to identify project areas likely to contain the types of sites most attractive to vandals, thus indicating priority areas for patrol and surveillance.

Planning

Perhaps one of the most valuable applications of modeling is in the area of planning. Planning for the management of cultural resources can take place during the development of land-use plans, environmental assessments, statewide or area-wide program plans, or site-specific plans. Models are especially suited to planning applications, because they focus on broad-scale, generalized trends, actions, or information. The main weakness of models, the inability to consistently produce detailed site-level specific statements, is usually not critical in a planning situation.

Modeling can help to plan how to reduce or anticipate adverse effects on cultural resources. For example, a model may predict the locations of sites that because of their complexity or their cultural or religious value to a Native American group may not be suited to data recovery. On the other hand, a model may predict the locations of sites that *are* suitable for data recovery of various kinds; estimates of potential costs and time needed for data recovery can then be derived by projecting site distributions in the planning area. Potential long-term and cumulative impacts to site pools by a proposed action can also be estimated, based on modeled populations of various site types. Modeling projections of so-called sensitivity areas have been widely used for planning purposes.

HOW CAN MODELS BE EVALUATED?

There are realistic limitations on the level of accuracy we can hope to achieve in locational modeling, owing primarily to the complex nature of the behavioral factors influencing site location. Models are simplified constructs of a complex universe that are seldom clearly right or wrong; rather, they are best viewed as being more or less useful. Often a model will excel in one application but fail in others. It is important for a manager to know what criteria of success are most important to the proposed application of the model, before embarking on a modeling project. While it is unrealistic to expect models to work with "perfect" predictive accuracy, it is not unrealistic to expect to know how well a particular model works and why. Indeed, this information is critical in deciding how the model should be used.

Several authors have discussed various criteria for model evaluation. In Chapter 2, Kohler presents an extremely useful discussion of inductive and deductive models, addressing their application, complexity, internal consistency, and precision. The appendix by Thoms carries this approach further, extensively comparing 22 models.

Undoubtedly the most important criterion to consider in evaluating a model is whether or not it has been tested. As noted earlier, untested models can be developed and formalized by using existing data, much of which contains biases. Simply because a model is formally stated, however, one should not assume that it has been tested or that its performance has been evaluated. Without testing or evaluation, a model is little more than a guess.

One of the main reasons for testing a model is to control for spurious or false correlations between site locations and the environment in a particular sample. Such correlations can be minimized by reducing chances for bias in the units selected for model testing, i.e., by avoiding artificial constraints and by selecting sample units randomly. Consider, for example, a sample inventory in which only sample units falling within 2 mi of a modern road have a chance of being selected. Analysis of site locations might reveal a marked correlation with geographical

variables that actually have more to do with road engineering requirements than with past human settlement preferences.

Establishing a control group to be used in model testing has been discussed in Chapter 8. This procedure is useful for model testing because it provides a background or baseline picture of the study area as a whole, against which model-generated statements can be evaluated. If, for example, it is noted that 90 percent of recorded sites are located within 5 mi of water, this observation could be very significant. If 90 percent of a control group of siteless locations in the project area were also found to be within 5 mi of water, however, the model would have told us nothing of significance about site locations.

Another reason for model testing is to determine the nature and strength of relationships that may have been discovered. Often several environmental phenomena occur together in nature (a relationship known as autocorrelation). Testing can help us to understand which of the co-occurring variables exerts the greatest influence on site location, and this information in turn permits us to evaluate the explanatory potential of the model. Perhaps the most obvious reason to test a model is to determine its overall accuracy rate. Accuracy rate and precision are generally inversely related. The more precise a model-like statement, the less accurate it is apt to be. This relationship is treated below in greater detail.

The procedures used in model testing have been treated extensively throughout this volume. Procedures for model validation and generalization are presented in Chapter 5. Various strategies for testing models based on existing data are presented in Chapter 7, along with techniques for integrating new data. In Chapter 8 the discussion covers several quantitative methods that can be applied to data collected through some form of probabilistic sample and carry with them some form of reliability measurement, such as confidence intervals or probabilities. The gain statistic is suggested as a useful measurement for comparing accuracy rates among models. Three rypes of testing procedures are described in order of increasing precision. Two, referred to as split sampling and the jackknife method, are based on testing the model against some portion of the original data used to develop the model. A third involves collecting new and independent data from the project area.

Several discussions of management concerns and model testing occur in this volume. In Chapter 3, Altschul distinguishes between wasteful errors (where a model predicts a site and none occurs) and gross errors (where a model predicts no sites and sites occur). In the latter case, the potential for inadvertent site destruction in many management applications is increased.

In Chapter 8, Kvamme discusses reduction of gross errors by adjusting the cutoff point of a model's decision boundary (a mathematical boundary), an approach that applies only to quantitative models. As an example of the relationship between gross and wasteful errors, perhaps a model permits us to say that 80 percent of the sites in a study area will be located on 50 percent of the land surface in that area. This represents a substantial reduction in the amount of land surface to be addressed further, but it also carries with it the potential for gross errors affecting 20 percent of

the sites. Using the same model, but adjusting the cutoff point, we may be able to say that 95 percent of the sites will be located on 70 percent of the land surface. This reduces the risk of gross errors, while increasing the possibility for wasteful errors. The implications of this discussion for management applications are significant. Reducing study area size by 30 percent would represent a substantial and desirable increase in project efficiency, especially if it could be accomplished with little or no risk to the resource.

In Chapter 3, Altschul cautions against passing the point of diminishing returns in model testing. This occurs when substantial increases in collection of new inventory data result in little increase in accuracy. There are many possible causes of this phenomenon, including the influence of such social factors as presence of large habitation sites, trade networks, and kinship groups, which override the influence of factors of the narural environment in determining site location and which are not addressed in the modeling effort.

An important consideration for evaluating models is their ability to take into account rare sites. These sites constitute a very small portion of the site population either by virtue of their own characteristics or by virtue of their location in relation to the environment. A site type can be rare without being impossible to model; most models do not address these sites, however, because their low numbers make most statistical techniques unusable.

The rare-site problem increases when sample inventories at low sampling rates are used to generate the data base for model development. When only a small percentage of the surface area is surveyed, the chances for discovering a rare site clearly are reduced. If *any* sites of a rare type are known in the study area, specialized inventory strategies can sometimes be devised to increase the potential for discovering more of these sites. If large village sites have been found only in riparian areas, for example, riparian areas could be sampled at a higher rate than other areas to increase chances for discovering this type, and compensation for the higher proportion of riparian areas surveyed in an otherwise random sample can be achieved during later analyses of the data.

Several other factors, any one of which could seriously affect a model's validity and usefulness, should be taken into account in evaluating a model. The manager should carefully analyze the appropriateness of all statistical procedures and analyses used in model development. Common problem areas include biases in the sampling procedures, failure to meet statistical assumptions about the data, and inappropriate use of environmental data. Often, for projects incorporating advanced statistics, the services of a professional mathematician will be needed.

Models should be evaluated for their completeness. Did they address changes in the environment through time? Are there biases in the sample design that might affect the reliability of the data? Also, the resolution of the model is important. If the management need is for statements specific to quarter-section parcels, broad zonal models may not be useful.

Inductive or correlative models have limited explanatory value because they do not account for observed correlations between independent and dependent variables. For example, empirical analysis may demonstrate that a certain type of site in a sample is always located within a limited distance of outcrops of a particular geologic formation. While this information may be very useful in certain contexts, it has not been demonstrated that the presence of outcrops actually *influenced* site locations. Independent evidence, such as the presence of specialized features or artifact types, is needed to support such an interpretation. This does not mean that the observed correlation is or is not valid; it means that we cannot explain why it occurred and thus we are no closer to an understanding of the broader cultural system that we are attempting to model. The utility of the model is limited to the observed study area. The need for independent testing to establish a noncoincidental relationship between the independent and dependent variables, in this case outcrops and site locations, is especially important because of the strong tendency for autocorrelation among environmental variables. Correlative models are useful because they direct these independent tests.

Field procedures are another factor to consider when modeling projects are being evaluated. For instance, the spacing of crew members and procedures for defining and recording sites can significantly affect the kinds of data that are available for analysis. Biases in field procedures should be explicitly stated in project reports, and their impact on the results of the modeling efforts should be evaluated.

Finally, the interpretability of the model is important. Is the model simple enough to be understood and explained in anthropological terms? Does it relate environmental and site variables to the everyday world? If not it may not be usable by future researchers in a cultural resource management context.

FUTURE DIRECTIONS

Predictive modeling holds much promise for cultural resource management in land-managing agencies, even though it is currently in a highly experimental and rapidly changing state. The information in this volume is not intended to limit or confine this development; rather, the intent is to crystallize issues and focus discussions on a common ground, to the benefit of both the agencies and the professional archaeological community.

At the present time, no major policy directives have been issued by a large land-managing agency concerning the development and use of models in cultural resource management programs. Many would argue that such directives would be premature. Many others would argue, however, that modeling has ceased to grow and contribute to our understanding in the way that it should because of a lack of focus and purpose in agency efforts. Altschul summarizes this concern in Chapter 3:

> Perhaps the most significant criticism that can be made about predictive modeling programs in most cultural resource management contexts is that there is no consensus as to the overall objective of these programs.

Current efforts are seen as diffuse and lacking in momentum and direction. Rather than working toward refining existing models or developing new types of information or methods, agencies sometimes develop new models that suffer from the same limitations as previous ones. Short-term goals are being pursued exclusively, perhaps because long-term goals have never been clearly defined or because incremental, long-term funding has never been available.

The information in this volume should help agencies to identify means of increasing the efficiency and effectiveness of their modeling efforts. The first, of course, is to develop the existing data base so that maximum use can be made of previously collected information. In addition, agencies need to augment the expertise of their staffs so that they will be able to evaluate and participate more fully in the modeling programs. This might involve specialized training courses in the evaluation and application of models and especially in the use of sampling techniques. Although advanced statistics will no doubt remain beyond the reach of the average staff person, some basic training courses in the types of models and their assumptions and requirements may be helpful. Only through this kind of staff development will agencies begin to use modeling effectively and creatively to direct and develop projects meeting specialized data needs. Only through this process will modeling be used as a long-term strategy, where it can be most effective and efficient.

There is a clear need to develop new ways to measure and define both dependent and independent variables. This can involve manipulation of tremendous amounts of information, for which remote sensing technology and geographic information systems are essential. Excellent and detailed discussions of these topics are provided by Ebert (Chapter 9) and Kvamme and Kohler (Chapter 10). Agencies should be aware of the potential contribution of geographic information systems to cultural resource modeling and make special efforts to ensure that the needs of the cultural resource program are met in the design of these systems. Because the potential contribution of a GIS is significant, consideration should be given to funding specialized research projects to explore possible applications of this new technology.

Finally, agencies need to focus on the development of explanatory theory. The kinds of information that can be obtained through traditional cultural resource surveys are limited. Surface observations made during the course of these surveys are based on "best guess" estimates of limited types of information. While this information is useful in the formulation of ideas and hypotheses about prehistoric societies, qualitatively different types of information are often needed to develop and test explanatory theories. This information, on topics such as diet, environmental exploitation patterns, technology, etc., can often only be collected through subsurface testing and excavation, accompanied by detailed laboratory analyses and studies, and through analysis of pertinent ethnographic, historical, and other nonarchaeological data. These approaches involve additional costs and for this reason are often not included in standard inventory approaches.

In order to further the development of explanatory theory and to increase the accuracy and usefulness of modeling efforts on a larger scale, agencies should seriously consider sponsoring research projects designed to measure complex social and economic parameters as they apply to the archaeological record. In Chapter 4 Ebert presents an excellent discussion of an innovative approach known as "distributional" archaeology. Here, traditional site types are seen as artificial constructs developed by archaeologists, which at best only poorly reflect behavioral systems. Analysis is focused on distributions of artifacts across a landscape as they relate to larger patterns of land use. Experimental work using this technique has taken place already in several land management contexts, and it appears to hold much promise for future advances in explanatory theory. Efforts such as these should not only serve to advance the state of predictive modeling, they should increase the efficiency and effectiveness of cultural resource management programs as well.

Chapter12

AN APPRAISAL

W. James Judge and Daniel W. Martin

In December of 1981 the Bureau of Land Management issued an instructional memorandum encouraging the development and use of predictive modeling in cultural resource management. Initial official interest in modeling by the bureau was in conjunction with the timely processing of "Applications for Permits to Drill" (APDs) for oil and gas. The oil and gas industry had recommended that the bureau initiate "regional reviews to identify areas of high and low probability for significant cultural resources, as a means for eliminating unnecessary surveys." The assumption was that "given an adequate data base, informed decisions can be made about where to concentrate additional identification and protection endeavors, to the exclusion of certain other areas" (Burford 1981).

The direction given by BLM headquarters at that time was as follows:

> States with heavy APD workloads are encouraged to consider developing predictive or sensitivity models for areas where it appears that cultural resource density and distribution lend themselves to the approach. Any such efforts should be directed primarily toward areas with high demand, where there is also an esisting basis for the expectation of a relatively low site population, regularity of site situation, similarity of site information potential, or other reasons for anticipating that the exercise will lead to a product that alleviates the cultural resource identification demands on BLM and industry, without creating an unacceptable risk to cultural resources [Burford 1981].

In attempting to implement the memorandum, resource managers found that predictive modeling was being employed in a wide variety of ways and that there was little mutually agreed-upon theory, method, or policy to guide the use of this technique. As a result, a proposal was developed by the BLM to fund a project that would address these issues. The project was approved and funded, resulting in the production of this volume.

The proposal established the following goals for the project:

1. to evaluate trends in the development and application of predictive modeling critically, using knowledge gained through past research efforts;

2. to explore the feasibility and practicality of predictive modeling for meeting management objectives;

3. to analyze and define the components of the model-building process, particularly with respect to cultural resource management;

4. to develop a set of standards for archaeological and environmental data to be used in modeling efforts; and

5. to provide BLM field officers with information on data collection for modeling purposes and statistical manipulations of those data.

The process by which the authors, editors, and advisory committee were selected and the lengthy course of peer and federal review to which the draft was subjected have been discussed in Chapter 1. In this chapter we will *(a)* evaluate the volume in regard to its success in achieving the goals outlined at the beginning of the project, *(b)* summarize the results of the peer review, and *(c)* discuss what we consider to be several important issues raised by this volume.

EVALUATION OF PROJECT GOALS

In general, the five goals presented in the initial project proposal were realized. The first, that of critically evaluating trends in the development and application of predictive modeling, is thoroughly addressed throughout the volume.

The second objective, that of determining the feasibility and practicality of predictive modeling as a useful technique for meeting federal management objectives, is addressed extensively in Chapter 11 and will also be discussed later in this chapter. We may note in passing, though, that to a certain extent apparent "success" in meeting this goal depends on how those federal management objectives are perceived. For some managers in 1983, predictive modeling was viewed as a technique that was going to rescue them from the burden of compliance with Section 106, permitting them to get by with minimal field survey and thus minimal expenditure of very scarce funds. To those individuals, the results of this volume may well be disappointing. To those who were looking for the satisfaction of more general, long-range objectives, the results will be received much more favorably.

The third objective, that of analyzing and defining the components of the model-building process as they apply to cultural resource management, is also addressed in detail in this volume. It is apparent that model-building is a very complex and time-consuming process. Nevertheless, there is freedom of choice as to how to proceed with modeling, and some ways of putting it all together may be more effective than others, depending on the situation and the needs. Again, Chapter 11 offers step-by-step considerations to guide modeling efforts for those with land managing responsibilities.

The fourth goal, to develop a set of standards for the archaeological and environmental data required to prepare predictive models, is somewhat more difficult to evaluate. In the literal sense, little in the way of a set of standards was developed by any of the authors. Their reluctance to provide a "cookbook" approach—which is implicit in the concept of standardization—is understandable,

given the variability in modeling approaches and management objectives, as well as regional physiographic and cultural differences. If, however, we consider "standards" to be a set of guidelines for data requirements in the model-building process, then the goal was met since the data requirements are standard in the sense that those agreed on as acceptable are presented in detail. For example, tolerable levels of error for data entry, choice of appropriate soil survey detail (e.g., Soil Surveys I, II, III), and appropriate cell-size choice for DEM (Digital Elevation Model) data are among the "standards" presented. Importantly, it is noted that each of the choices made must be tailored to a specific objective and phase of the modeling process and to specific regional circumstances.

In one sense, groundwork for development of more standardized data is provided in this report. Perhaps the best way to establish such standards would be to develop them from data used in actual field and management applications. Standards developed in this way would thus be based on actual management successes and would minimize the possibility of error.

With respect to the final objective, that of recommending types of field inventory data to be collected and of developing specific procedures for field office use, only the initial part of this goal has been met in detail: recommendations regarding field inventory data are found throughout the volume. The second part is left quite open, again because of our reluctance to provide a cookbook approach, and also to enable field offices to pick and choose among techniques themselves so that local management needs are addressed by the most efficient means.

In our view, then, the objectives of the project were effectively met, particularly when one considers the complexity of the subject matter, and the absence of a well-developed body of theory and method for predictive modeling when the goals were established.

AN APPRAISAL OF THE REVIEW COMMENTS

This volume benefited from extensive peer review. The invitation to review was extended to numerous organizations in order to create a document that represented participation from a broad spectrum of the professional archaeological community. Comments were received from the following organizations: Bureau of Land Management offices, State Historic Preservation offices, the National Park Service, the Department of the Army, the Bureau of Indian Affairs, the Advisory Council on Historic Preservation, the Bureau of Reclamation, the Forest Service, the Soil Conservation Service, the Society for American Archaeology, and a number of universities. The responses provided substantive comments on theoretical, methodological, technical, management, procedural, legal, and regulatory issues presented in the draft version. Even the most critical reviewers felt that the volume was an important contribution and should be published.

Many of the comments suggested that the dichotomy between correlative and explanatory modeling was artificial and that the importance of explanatory models

was over-emphasized (especially as being superior to correlative models). Some felt that the dichotomy between the kinds of models was useful primarily in a heuristic sense, while others supported the research commitment to explanatory modeling but felt the important role of the correlative approach in the development of predictive models should be acknowledged. Some comments noted that in the normal scientific process such contrasted approaches are actually complementary, but that the empirical search for patterns may well precede the quest for explanation.

A majority of the reviewers felt that the report was too negative about the potential of modeling in CRM contexts. Most of the federal reviewers felt that the early chapters were unnecessarily "academic" or pedantic, and that more practical advice was needed (Chapter 11 was not available with the review draft). Some of the polemic regarding distributional archaeology (Chapter 4), for instance, and the extended debate regarding inductive and deductive issues were felt to be of little value by this group of reviewers.

Archaeologists with management responsibilities feared that the suggested potential of predictive modeling was too limited. They were looking for practical methods to provide better information about cultural resources in order to make realistic recommendations to management. Archaeologists without management responsibilities appeared to fear that the technology, if allowed to go unchecked, would be applied by the government in an irresponsible manner. In this vein, federal reviewers felt that the orientation of the volume appeared to be toward archaeologists without management responsibilities.

All in all, the peer review comments, which themselves comprise hundreds of pages, proved to be extremely helpful in guiding the development of the final volume in a direction most useful to the diversity of the anticipated audience.

THE ISSUES RAISED

A number of key issues have been raised in this volume regarding the relationship between an emergent technology based largely in theory and practical everyday management needs. Here we will summarize four of the issues that we feel are extremely important to the topic of predictive modeling for both archaeological research and cultural resource management.

The first issue is that of the complexity of the process; modeling past human activities is not a simple task. Humans, fortunately, do not behave mechanistically, and thus generalizations about their behavior are difficult to derive and can never be completely accurate. The relationships among humans, their activities, and past landscapes are very complex to begin with, and this complexity is increased by subsequent changes in those landscapes, by a depositional record that is both incomplete and complex, and by the difficulty of the quantitative methods that one must employ to model these relationships—methods that are frequently beyond the expertise of those who wish to use them. Modeling is a tool, but it is by no means

a simple tool and it is not a panacea. As a complex tool, its uses are limited, and it requires expertise to implement correctly. As with any tool, modeling can be abused, and the value of the results diminishes accordingly. Used properly, however, modeling can be of inestimable value to both the manager and the research archaeologist. This volume, we feel, presents the complexity of the modeling process well, and Chapter 11 details its appropriate uses in the management context.

The second issue raised is that of the role of predictive modeling in the compliance process, that is, in efforts to comply with Sections 106 and 110 of the National Historic Preservation Act. This, of course, was one of the key concerns that stimulated the project in the first place. Managers were almost desperately seeking some way to address compliance problems in a cost-effective manner that would also protect the resources. Archaeologists may have felt that cost-effectiveness was taking precedence over resource protection, but many managers saw the situation differently. Shortly after the release of the BLM instructional memorandum noted at the beginning of this chapter, a project was proposed by BLM staff that was to use

> statistical discriminant analysis techniques to develop a model to predict the probability of cultural resource occurrence from environmental parameters and evaluate the utility of this methodology as a tool in cultural resource assessment on potential oil shale and coal lease areas. . . . Once the model is developed and tested it can be turned over to the District or Area Office Archeologist where it can be used operationally to predict the probability of site occurrence on rights-of-way applications, access corridors and drill pad clearances. If in this stage high probabilities are present, the corridor could be moved to a lower probability zone. In other cases, the probability could be used to ease the requirement to have a site visit prior to clearance [Garratt 1982].

Clearly, managers were having problems with the compliance process, and expectations that predictive modeling would solve or lessen those problems were high.

In Chapter 11, Kincaid points out that Section 106 compliance decisions are made on a case-by-case basis through the consultation process, and that there are no set criteria for determining appropriate inventory and evaluation strategies apart from such consultation. In brief, there can be no "cookbook" approach to the role of modeling in that process. We can, however, summarize the value of modeling in the inventory process in general, whether it be for research or management purposes.

Predictive modeling of archaeological site locations can never be a complete substitute for actual field inventory (intensive survey). As noted above, not only is human behavior too complex to permit this kind of modeling accuracy, but too many variables have intervened between the time that the behavior took place and the present to allow us to achieve through modeling the accuracy available with field inventory. For this reason, it is unlikely that predictive modeling could, in the foreseeable future, be sufficiently accurate to satisfy the identification requirements

in 36CFR800.4 (the implementing regulations for Section 106 of the National Historic Preservation Act; see also Secretary of the Interior's Standards and Guidelines, Federal Register 48(190):44721–44723). By the same token, predictive modeling is unlikely to satisfy the needs of a research archeologist whose research design requires accuracy at a similar level.

Modeling can, however, provide research archaeologists with estimates of probable site densities in unsurveyed areas, and this same capability is of great potential benefit to the manager. As noted in Chapter 11, the role of modeling in the planning process is perhaps its most valuable contribution. In the short term, for example, the ability of models to project areas of low site density or to indicate probable locations of sites not suited for data reeovery can be extremely helpful to the manager, not as a substitute for inventory but as an aid in designing cost-effective inventory.

Modeling's greatest strengths, however, lie in its contributions to the long-term planning process. It is here that models developed with resource planning, interpretation, and evaluation in mind can be of tremendous value to the establishment of management priorities and to the integration of cultural resource management with other resource management responsibilities. Further, such model-based management can facilitate research, quite apart from the preservation and protective responsibilities of the manager. Since a fundamental purpose of cultural resource preservation is to maintain the scientific potential of the resource, that is, to preserve its information content, modeling as a component of long-range planning is of particular value to managers and researchers alike.

The rhird major issue raised in the volume has to do with the theoretical basis of predictive modeling. Certainly the volume provides a critical summary and evaluation of current perceptions about the relationship between modeling and theory. Aspects of theory dealt with include examination of the systemic, archaeological, and analytic contexts, as well as site formation processes. Normative vs processual theoretical approaches as they relate to modeling efforts are also detailed.

The most fundamental theoretical issue to emerge, however, is that of the dichotomy between correlative and explanatory models. This dichotomy arises from the contrast between inductive and deductive logic, although the terms *deductive* and *explanatory* and the terms *inductive* and *correlative* are not synonymous. Technically, models themselves are either explanatory or correlative; the terms deductive and inductive refer to how the models are derived and to the kinds of arguments involved in their implementation. Correlative models tend to be inductively derived (but not exclusively so), and explanatory models should contain arguments of both types.

In Chapter 2 this theoretical dichotomy is discussed with respect to the various contexts in which archaeological investigations are carried out.

> The challenge for inductive models is to build the bridge to the systemic context by making the analytic methods (including discovery) as "transparent" (non-bias-making) as possible and by controlling for the effects of depositional and postdepositional processes in the archaeological context.
>
> Deductive models, on the other hand, begin with some theory predicting human behavior, in the systemic context. The challenge for deductive models is to build the bridge to the analytic context, which is where the outputs of the system can be observed. This bridge-building—whether from the systemic to the analytic context or vice versa—is referred to as explanation. . . . Explanatory models . . . are inherently neither inductive nor deductive. Instead, they are models that attempt to build the bridge between the dynamics of the living system and its observed outputs [Kohler, Chapter 2].

As noted in Chapter 11, the contrast between correlative-inductive and explanatory-deductive modeling becomes somewhat blurred in field modeling applications. In actual practice, correlative models are generally easier to develop and in specific situations may be more accurate in their predictive potential. These models are criticized, however, for their lack of ability to explain the phenomena predicted. Archaeologists are concerned about the explanation of past human behavior, and there is general agreement that we should not be satisfied with only the demonstration of correlations, but that we must also provide explanations for those correlations. Even if it is acknowledged that archaeologists consider explanation to be the goal of modeling, however, a fundamental question still remains: how necessary is such explanation to the actual everyday management of cultural resources? This, in itself, is a key issue raised by this volume.

As noted above, archaeological resources are most often preserved for their information content. There is no question that the inherent information can best be extracted through the explanatory process, and correlative models, because they are derived inductively, cannot contribute as much to the extraction of this information as models with a consciously explanatory orientation. But this is not the central question in cultural resource management. In that context we must ask, what is the best technique to preserve the resource? What is the most cost-effective means to achieve preservation, and to what extent is explanation necessary for effective management? By "preservation" here, we refer to the full complement of tasks involved in resource management, including discovery, recording, evaluation, conservation, and protection. There are no simple answers, but we may offer some comments.

Basically the issue is this: should the manager select a correlative model, which is easier to design, takes less time to develop, and is initially more accurate, or should he or she plan to use an explanatory model, which is more complex and difficult to develop and may not be as accurate a predictor? At first glance, the answer would seem to be simple: go with the correlative model, and let archaeologists with research interests develop their own explanatory models at some time in the future. In that way, the resource will have been protected in a cost-effective manner. After all, management is under no legal obligation to provide explanation as part of the preservation process.

Yet the decision is not that straightforward. Correlative models are not immediately "transferable," that is, when developed for one geographic location, they do not necessarily work in another; there is no logical reason they should. The question then is whether it is more cost-effective to redevelop (or at least refine and reaffirm) the correlative model for use in a new area or to develop the explanatory model in the first place, since the latter would be applicable in a variety of areas and would address other management needs (interpretation, evaluation) at the same time. Ultimately, this question can only be resolved on a case-by-case basis where all the variables to be considered can be evaluated properly. But certainly prior to investing time and funds in the development of an explnatory model, the manager must determine whether it actually would be as easily transferable as claimed and whether it will be accurate enough to satisfy resource preservation and protection requirements. We feel that research and management archaeologists alike would agree that, if one has the time and funding, explanatory models will be more generally productive in the long term, and thus ultimately more cost-effective. But such decisions must be made for each specific instance by managers, employing the best information possible at the time.

One further aspect of the dichotomy between the two types of models is the supposition that explanatory models may serve management better in the process of site evaluation. There is little question that the determination of the significance of a site, or class of sites, may be enhanced by the deductive process integral to explanatory model development. Yet at times significance may have to be determined on the basis of the resource's potential, rather than the demonstrated contribution of information. This is true in archaeology, where sites frequently cannot be excavated, and thus the information content cannot be fully demonstrated through deductive testing. In such cases, the potential significance is assessed from surface indications, and at this level of evaluation, correlative models may be as effective as their explanatory counterparts in indicating a resource's potential contribution to scientific knowledge. Again, the cost-effectiveness of redeveloping correlative models for use in other areas may be the key decision that managers have to make.

A fourth issue raised in this volume was that of the technology and expertise necessary to implement modeling effectively. Sophisticated hardware and software capabilities are requisite, as well as well-trained and informed individuals at all managerial and support levels.

For example, it has become clear that successful application of certain models may require the use of a geographic information system (GIS). The quantity as well as the quality of analyses necessary require automated spatial analysis of data. Remote sensing techniques provide a source of data for GIS analysis. The availability of multispectral, high-resolution digital imagery opens up exciting possibilities for pattern recognition techniques presented in this volume. The dramatic leap to 10 m resolution by the SPOT satellite is only the beginning; far more detailed resolution will be available in the future. The scale of measurement of the instrument has been one limiting factor, along with limited processing capabilities for

gigabytes of data. These technologies are improving, and the speed of this improvement provides an insight to the level of refinement we may expect from the modeling process in the future. The basic statistical, modeling, and pattern recognition theories, methods, and techniques presented herein provide the foundations upon which to build powerful new instruments of measurement and analysis. At present too few people in management and support positions have the requisite skills in geographic information systems (Burrough 1986), statistics, remote sensing, and modeling to exploit the technology available currently, let alone develop future applications.

Another problem is that of obtaining access to the most capable systems and to adequate data bases. Access to such systems with diverse data themes and regular data maintenance is most readily available to persons who work for, or have some formal connection with, large land-managing agencies. Such systems require an organizational support structure difficult to justify for single-purpose analysis. Large land-managing agencies are supporting such systems on the basis of their utility to overall land management analysis. Included in such support is providing quality software and hardware, software development, management, and various levels of staff skills, training, and technical assistance.

These are some, but by no means all, of the issues raised in this volume that we feel are extremely important to both research and resource management as they relate to predictive modeling. The issues that have not been summarized here may have equal significance in particular modeling applications. One of the purposes of this volume has been to bring a wide range of issues in the domain of predictive modeling to the fore.

CONCLUSIONS

Predictive modeling can clearly be a worthwhile component of cultural resource management, if for no other reason than that it injects rigor into the management process and serves to integrate management with archaeological research. The process of modeling and the preparation and development of models are extremely valuable assets to management, regardless of the ultimate "success" of the models.

After a thorough review of predictive modeling, this volume reaches some conclusions that contradict past attitudes and expectations held by land-managing agencies. The Bureau of Land Management's proposal noted previously (Garratt 1982), for example, dealt with only a part of an overall process. We have learned that the application of "statistical discriminant analysis techniques" to environmental variables is not sufficient to develop a usable model. Certainly, the proposal made the process sound too easy and neglected much detail. We have learned that we must be sensitive to the facts and theories of site formation processes, and that it is necessary to incorporate theory from anthropology, archaeology, and other social scientific disciplines because site distribution is a reflection of human behavior

interacting with physical phenomena in an ecosystem. Again, calling attention to the complexity of predictive models and the modeling process is an important contribution of this volume.

Further, we have learned that modeling is a cyclical process of ongoing refinement, rather than a one-time event, and thus models cannot be developed by outsiders and then simply "turned over" to agency field office archaeologists for "application." For many reasons the field archaeologists and managers need to be full participants in the modeling process. We can conclude that predictive modeling, as defined and developed herein, is potentially the most cost-effective way to combine sound management practices with valuable research programs. Both are necessary ingredients for cultural resource preservation and interpretation in this country.

It may well be that the most cost-effective and appropriate manner for managers to implement the techniques discussed in this volume would be to focus on the development of correlative models initially and then work toward refining their accuracy. This will demonstrate the potential of modeling and its effectiveness as a tool for cultural resource management. But the correlative-inductive approach should never be considered an end in itself. Instead these initial models should be specifically designed as integral components of the deductive approach to model development and as parts of the long-range planning process necessary to achieve the full potential of predictive modeling in resource management through ultimate reliance on explanatory models.

REFERENCES CITED

Burford, Robert J.
1981 Cultural Resource Identification Requirements for Onshore Oil and Gas Drilling Permits, Instruction Memorandum No. 82-162, 8143(341), December 23, 1981. Bureau of Land Management, Washington, D.C.

Burrough, P. A.
1986 *Principles of Geographical Information Systems for Land Resources Assessment.* Oxford University Press, Oxford.

Garratt, Michael W.
1982 Cultural Resource Assessment Model Development, Development Project Proposal, Form 1681-1. Bureau of Land Management, Service Center (D-440), Denver Federal Center, Denver.

Appendix

A SURVEY OF PREDICTIVE LOCATIONAL MODELS: EXAMPLES FROM THE LATE 1970S AND EARLY 1980S

Alston V. Thoms

The purposes of this appendix are to expose the reader to a range of projects that have developed predictive models and to provide succinct comparative summaries of these projects. A variety of geographic areas, archaeological manifestations, and modeling approaches are represented. Twenty-two projects were judgmentally selected from more than 100 reports. The longer list was not exhaustive; it reflected the interests of the authors of this volume and was generated by combining lists of references provided by the authors and by the project advisory team.

The projects summarized here represent a range of approaches and are *not* limited to the best or most successful examples; indeed, *best* and *most successful* are terms that would be difficult to define in a manner acceptable to all readers. Projects employing state-of-the-art approaches and some earlier examples of predictive models are included, as are examples of the less successful approaches. Information about the characteristics of what may be unsuccessful predictive models can be useful in providing the reader with a broad data base against which the usefulness of predictive models under a wide range of conditions may be evaluated.

Among the 22 projects summarized here are studies from many portions of the United States (Figure A.1), from projects in Delaware (Custer et al. 1984) and Georgia (Kohler et al. 1980) to those in Washington (Mierendorf et al. 1981) and Alaska (Ebert and Brown 1981). The emphasis, however, is clearly on the western states (e.g., Bradley et al. 1984). Included in the sample are models that predict the distribution of sites that are visible on the surface (Larralde and Chandler 1981), of sites that are deeply buried in flood basins (Muto and Gunn 1980), and of inundated sites on the continental shelf (Barber and Roberts 1979). Predictions of site distributions are made for relatively undisturbed areas of the Great Basin (Tipps 1984) and for highly developed areas along the eastern seaboard (Hasenstab 1983). There are models for predicting the density of sites in areas occupied by mobile, montane hunters and gatherers (Jermann and Aaberg 1976), and models concerned with more sedentary Anasazi farmers (Woodward-Clyde Consultants 1978). Much of the time span of human occupation in North America is represented by these models. There are predictions for the locations of sites occupied by the earliest inhabitants of the

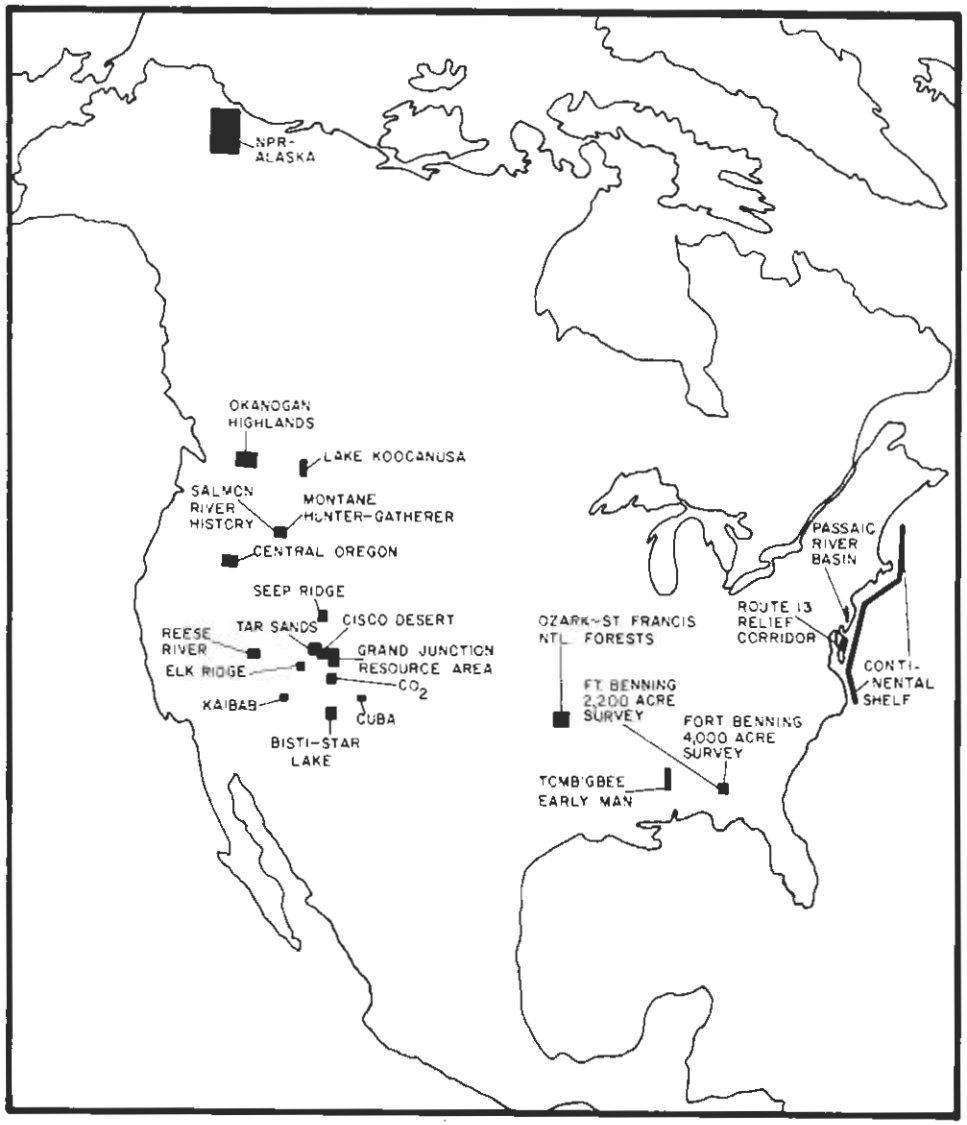


Figure A.1. Map of North America showing approximate locations of predictive modeling projects discussed in the Appendix.

Ozark highlands of Arkansas (Sabo et al. 1982) and predictions for the locations of recent Euroamerican ranches in the Salmon River Mountains of Idaho (Rossillon 1981).

The project summaries encompass deductively derived models (Thomas 1973) and inductively derived models (DeBloois 1975), including a deductive economic-decision-making model that predicts proportional use of the landscape (Hackenberger 1984) and an inductive landform-analysis model designed to predict the general location of significant sites (Wildesen 1984). Some of the models can be

tested with future survey data (Kemrer 1982), and other projects were developed as tests of existing predictive models (Thomas et al. 1983). Finally, the selected sample includes predictive models made by simple extrapolation from known to estimated site densities in large environmental zones (Plog 1983a) and very complex models developed using multivariate statistics and geographic information systems to generate probability estimates for site presence/absence in areas covering less than 1 ha (Kvamme 1983).

Once the selection of project reports to be summarized had been made, it was necessary to develop a list of attributes or variables that could be monitored for each report. The attributes monitored are *(a)* project location and size, *(b)* inventory method, *(c)* analytical techniques, *(d)* the nature of the model used or developed, and *(e)* the success of modeling efforts. The evaluation of each project also includes a discussion of other relevant topics introduced elsewhere in this volume. Toward these ends, the reports were examined in some detail. What might be called a "mental regression analysis" was performed to identify those variables that could be monitored with reasonable consistency and related to the topics discussed (and to the terminology employed) in the various chapters of this volume. On the whole, the terminology used here corresponds most closely with that utilized by Kohler in Chapter 2.

The results of this survey of project reports are presented in two parts. The first part includes detailed information presented in a series of tables designed to facilitate comparisons of the various approaches. Summaries of each modeling project are presented in the second part, along with a few brief comments about the approaches used. Comments focus on the relationship between modeling objectives and results, as well as on innovative aspects of the methods employed. The overall discussion ends with some general observations about the nature of predictive modeling as represented primarily by the selected sample of project reports. Some of the comments are particularistic because they refer to a given aspect of a specific project. Other comments about a given project are made because that project is characteristic of a general approach to predictive modeling.

TABULATED SURVEY RESULTS

Descriptive and evaluative information about the reviewed projects is summarized in tabular form. Table A.1 provides information on general characteristics of each model—location, type of model (inductive or deductive), objectives, claimed accuracy (high, low, or percentage estimates), mode of presentation (tables, maps, charts), and verification approach (how the model was tested). This table also includes a general assessment or evaluation of each model. The evaluation criteria—*falsifiability* (can the model be disproved?), *consistency* (is it mathematically and logically sound?), *simplicity* (is it parsimonious?), and *generalizability* (can it be applied to other study areas and to human behavior in general?)—are essentially the criteria defined by Kohler in Chapter 2. An assessment is also made as to how

thoroughly the environmental and cultural data were evaluated before they were used in the model. This assessment includes such questions as whether there was an effort to reduce redundancy, whether the reliability of map-based information was discussed, or whether other statistical techniques were considered. Results of this systematically judgmental assessment are presented as scores on a scale from 1 (lowest) to 5 (highest).

Table A.2 characterizes the models in terms of their general data bases and predictions. It presents information regarding the kind of sampling procedure used, the number of sites or cells included, and the size of the cells, transects, or grid units used to subdivide the sample. Levels of measurement (nominal, ordinal, interval, or ratio) used to define or describe environmental variables are also listed, as is the nature of the predicted resources (site type, site density, or site presence). The manner in which the survey area was classified into landforms or environmental types and into site density zones or site present/absent units is also summarized. The spatial resolution (e.g., block areas, landforms, grid units of various sizes) of the predictions and the nature of the predictions (e.g., site density, site presence, site significance, or site type) are characterized under the heading "Resolution of Predictions." An evaluation of the thoroughness of the procedural discussions in the report is presented as a score on a scale from 1 (lowest) to 5 (highest).

Information related to the environmental variables used in the models is presented in Table A.3. The listed physiographic divisions within which the projects are located follows Hunt's (1974) classification. Major types of contemporary land use are also listed, as is the size of the project or study area (i.e., the extent of the spatial population for which predictions are made). Environmental variables used to classify or to subdivide the project area (e.g., landform type, soil type, distance to water, elevation, and slope) are listed, as is the source of that information (e.g., various kinds of maps, field observation, and literature search). The modeling projects are rated from 1 to 5 assessing *(a)* the degree to which changing paleoenvironmental settings are considered and *(b)* the degree to which the effect of various depositional environments on the discovery of cultural resources and/or on our understanding of past human behavior is taken into account. The same scale of ranking is used to assess the level of discussion about the ecosystems within which humans operated. In other words, the scale provides a comparative measure of how well the investigators discuss the spatial and temporal distribution of food resources that may have been used by past groups of people.

Cultural variables used in the modeling projects (e.g., site type, site size, artifact/feature types, or simply site location or presence/absence) are summarized in Table A.4. The culture area designation follows Driver's (1961) scheme. Terminology used for known and predicted site types usually is taken from the referenced report. The sources of information about these cultural variables are also tabulated. The models are assessed on a scale from 1 to 5 according to the level of consideration given to understanding the human land-use systems represented by the debris on or in the ground.

Table A.5 characterizes the nature and results of field investigations conducted to develop or test the models. In some cases fieldwork was not part of the modeling project; rather, existing survey data were used to build and/or test the models. For those projects for which new data were collected, information is provided regarding how the field data were used, the size of the survey area, and the general methods used to discover and/or record the resources. Some of the results of the fieldwork—number, types, and densities of sites discovered in the survey area—are tabulated. The general nature of the fieldwork is assessed by evaluating the reports (again on a scale from 1 to 5) according to the thoroughness of the discussion of constraints and limitations imposed by field methods. For example, is there a discussion of the kinds of sites that potentially remained undetected when subsurface deposits were not exposed (e.g., by clearing of duff or leaves, digging of test pits, or cleaning of existing cutbanks)? Did survey strategies result in the detection of the full range of known or theoretically expected site types? What were the effects of excluding areas from the survey or of arbitrarily distinguishing between sites and isolated finds on the basis of artifact density?

Project reports are listed in chronological order in the tables and in the following summaries in order to afford the reader an opportunity to assess developmental trends. They span the time period from 1973 to 1984; 15 of the 22 were published after 1980. Reports that were published or printed in the same year are listed in alphabetical order.

SYNOPSIS OF SURVEY RESULTS

The summaries presented in this section provide a brief synopsis of modeling components of the 22 project reports. This information is intended to fill in some of the gaps in the tabular summaries and to provide coherent descriptive statements for each model. Additional information is also provided about the institutional affiliation of the investigators and the funding agency for each modeling project. Attention is drawn to any special qualities or potentially undesirable aspects of the models. The concluding paragraph in each synopsis is essentially a narrative assessment of how well the modeling project achieved its stated or implied goals.

Reese River Ecological Project "An Empirical Test for Steward's Model of Great Basin Settlement Patterns." David Hurst Thomas. *American Antiquity* 38:155–176. 1973

The Reese River Ecological Project was conducted by the American Museum of Natural History and funded, in part, by the National Science Foundation and the University of California (Thomas 1973). It is one of the few research projects, as opposed to cultural resource management projects, selected for summarization.

TABLE A.1.

Summary of general characteristics of the selected predictive model projects

Project Location	*Type of Model*	*Model Objectives and Applications*	*Claimed Level of Accuracy*	*Model Presentation*	*Approach to Verification*	*General Assessment (Scale 1:5)*				
						Data Evaluation	*Falsifiability*	*Consistency*	*Simplicity*	*Generalizability*
Reese River Ecological Project (Thomas 1973)										
Central Nevada; Lander and Nye counties	Deductive logic in a systemic context	To test Steward's Great Basin Shoshonean subsistence-pattern model using archaeological data	Overall, 86% of quantitative predictions were verified (p. 171)	Tables, charts, text maps with a 500 by 500 m grid; 1:24,000 and 1:62,000 scales	New field data used to test expectations about specific artifact densities and frequencies by zone	3	4	4	3	4
Elk Ridge Project (DeBloois 1975)										
Southeast Utah; San Juan County	Inductive logic in an analytic context	To design a predictive sampling strategy as an interim step in the total inventory of archaeological resources	Sample surveys may "prove to give better results than a 100% survey done poorly" (p. 127)	Text, graphs, mathematical model, maps (1:24,000) with various grid unit sizes	Sample at different levels to determine which percentage provides the most reliable and economic results; chi-square test of selected site characteristics	2	3	3	4	3
Lake Koocanusa Project (Jermann and Aaberg 1976)										
Northwest Montana; Lincoln County	Inductive logic in an analytic context	To estimate the total number of archaeological sites within the reservoir; to provide planning information	NRO*, authors stress preliminary nature of findings and that results are best applied to upper terraces (pp. 2, 73)	Tables for predicted site density per unit area within sampling strata delineated on 1:24,000 and 1:9600 topographic maps	Untested, but subject to testing with new survey data	2	4	3	4	3
CO_2 Project (Woodward-Clyde Consultants 1978; James et al. 1983)										
Southwest Colorado; Montezuma and LaPlata counties	Inductive logic in an analytic context	To identify areas likely to contain sites and to differentiate among them on the basis of significance; to develop a graphic planing tool to provide a basis for improving well field design by minimizing potential impacts	About 80% of the variation in data is explained by regression relationships (pp. 6, 20)	Output in tabular, numerical, graphic form; final product is a "plan dimension cell plot" on 1:24,000 topographic maps	"Visits" to 140 cells (ca. 122 by 122 m) totaling about 1.5 ha	2	2	3	3	2
Continental Shelf Project (Barber and Roberts 1979)										
Coastal and Continental Shelf portions	Two approaches; inductive logic	To predict where sites will be found on the shelf and what types of sites can be	The DEductive model is suggestive, moderately sophisticated, reliable	DE: tables and maps (1:125,000); IN: maps	Untested, but subject to testing with new survey data	IN:4 DE:4	IN:4 DE:3	IN:3 DE:3	IN:3 DE:4	IN:3 DE:4

of Maine, New Hampshire, Massachusetts, Rhode Island, Connecticut, New York, New Jersey, Delaware, Maryland, Virginia and North Carolina	in an analytic context and deductive logic in a systemic context	expected; for planning purposes	(p. 351); the INductive model is "basically sound" (p. 330)							
Fort Benning 4000-Acre Project (Kohler et al. 1980)										
West-central Georgia (Muscogee and Chattahoochee counties); East-central Alabama (Russell County)	Inductive logic in an analytic context	To identify variables that determine site location and define site patterning; for use in interim management and planning of additional survey work	Probabilities of finding sites in zones are stated, but model is not tested; no claim for accuracy (p. 79)	Tables, charts, maps at scale of 1:25,000	Untested, but subject to testing with new survey data	3	4	4	4	3
Tombigbee Early Man Project (Muto and Gunn 1980)										
Northwest and east-central Mississippi and west-central Alabama; numerous counties along Tombigbee River	Basically inductive logic in an analytic context but with some deductive systemic elements	To reconstruct the late Quaternary environment so that the variables of culture change and continuity are identified and applied to a predictive model for locating Early Man sites	Of the 31 randomly selected locations, 16 or 51% were archaeological sites (pp. 5-284)	Tables, and more than 600 point-specific, probable site locations plotted on 1:24,000 or 1:62,500 topographic maps	"Empirical site locational model" tested with new field data to develop "initial predictive model" to be tested with future fieldwork	3	3	3	2	3
NPR-Alaska Project (Ebert and Brown 1981)										
North-central Alaska	Inductive logic in an analytic context	To use remote sensing methods to correlate environmental settings with known site locations and to contribute to the accuracy and cost efficiency of the cultural resources assessment in unsurveyed areas	Approach yields useful information in determining "upper limits" of site densities and occurrences of other site characteristics (p. 381)	Tables and 1:250,000 topographic maps with plot of known sites and simplified version of identified strata/site-density (ecologic) cover types and transitional areas	Untested, subject to testing with new survey data	3	3	3	4	3

*NRO: Necessary information Not Readily Obtained from the referenced source

(continued)

TABLE A.1. (Continued)

Summary of general characteristics of the selected predictive model projects

Project Location	*Type of Model*	*Model Objectives and Applications*	*Claimed Level of Accuracy*	*Model Presentation*	*Approach to Verification*	*General Assessment (Scale 1:5)*				
						Data Evaluation	*Falsifiability*	*Consistency*	*Simplicity*	*Generalizability*
Seep Ridge Project (Larralde and Chandler 1981)										
Northeast Utah; Uintah County	Inductive logic in an analytic context	To derive a formula to define the probability of site occurrence at any point; to define suspected extremely low site density areas for management purposes so that project-by-project clearances may not be necessary	With the discriminant function 97.1% reclassification accuracy for site/nonsite dichotomy	For sensitivity model, zones are plotted on 1:126,720 maps; the discriminant function model is presented in text with the formula	Discriminant function model for 34 Seep Ridge sites tested with data on 98 other sites, then 98 sites used to develop new model tested with 34 sites	4	3	3	3	3
Okanogan Highlands Project (Mierendorf et al. 1981)										
North-central Washington; Okanogan, Ferry, Stevens, Chelan, Grant, and Lincoln counties	Deductive logic in a systemic context, but with major inductive/analytic elements	To predict the location, density, and sensitivity (to construction impacts) of archaeological resources within a large and relatively unsurveyed area; to provide a long range planning document	Predictive model is "overly simplified approximation of cultural-ecological relationships"; starting point for prehistoric land use investigations (p. 98)	For predictive model, seasonal land-use zones (with identified site types) are plotted on a ca. 1:800,000-scale map; other maps of same scale illustrate zones according to significance values, relative site density/projected impacts, and composite of these for sensitivity model	Untested but could be tested with new survey data or statistical analysis of existing site file data	3	2	4	4	4
Salmon River History Project (Rossillon 1981)										
Central Idaho; Custer, Idaho, Lemhi, and Valley counties	Deductive logic in a systemic context	To predict livestock-raising site locations from the late nineteenth century as a means of facilitating site discovery and as a model for future testing	Model may predict general areas of winter and summer cattle- and sheep-raising sites (p. 70)	Flow charts and maps (ca. 1:833,333) illustrating seasonal range "palatability" and "predicted" hay production for cattle and sheep within units ca. 4828 by 4828 m (ca. 23.3 km²)	Untested but could be tested with results of detailed archival searches for farmstead locations, utilized meadows, etc.	3	2	3	4	2

Bisti-Star Lake Project (Kemrer 1982)										
Northwest New Mexico; San Juan and McKinley counties	Inductive logic in an analytic context	To develop, test, and refine "methods which formally predict" archaeological component densities (sites by time periods) across study area, most of which is unsurveyed	Expressed as explained variance and ranging from 46% to 80% depending on site type and version of model	Tables and maps (ca. 1:125,000) showing predicted site type in 2000 by 2000 m grid units	Model was constructed using site file information from adjacent areas and tested/refined using sample survey data from study area	4	2	3	2	3
Ozark-St. Francis National Forest Project (Sabo et al. 1982)										
Northwest and east-central Arkansas; Baxter, Benton, Conway, Crawford, Franklin, Johnson, Logan, Madison, Marion, Newton, Pope, Searcy, Stone, VanBuren, Washington, and Yell counties	A deductive model in a systemic context is developed using data from the literature and then assessed by comparing it to an inductive model in an analytic context	To assess potential nature and distribution of prehistoric and historical sites in unsurveyed areas so the predicted cultural resources can be incorporated effectively into multiple resource management plans	Results of statistical analysis are generally in agreement with expectations derived from prehistoric and historical models (p. 184)	Tables and overlays for maps (1:24,000) that depict site likelihood zones (by general topographic settings)	The deductive model is tested using existing site file data in a series of Q-mode cluster analyses that generate the inductive model	4	2	3	4	4
Passaic River Basin Project (Hasenstab 1983)										
Northeast New Jersey; Morris, Essex, Union, Somerset, Hudson, and Passaic counties	Prehistoric model uses inductive logic in an analytic context; historical model is revision of existing model using inductive logic in an analytic context	To estimate the quantities of cultural materials likely to be impacted and to define those areas with high probability of site occurrence as means of providing a "tool for project planners" in carrying out future cultural resource management measures	"Favorable from an academic point of view" but less desirable from a management view; models failed to "pare down the overall high sensitivity strata" (p. 157)	Tables, charts, and computer-generated maps (1:24,000) that illustrate aspects of model, including classification of each area according to its general likelihood of having prehistoric and historical sites	Historical and prehistoric models devised from GIS/univariate analysis were tested with new survey data and revised accordingly to yield "field-derived site encounter rates" for the sensitivity strata	3	4	3	2	2
Grand Junction Resource Area Project (Kvamme 1983)										
West-central Colorado; Garfield and Mesa counties	Inductive logic in an analytic context	To develop quantitative models for predicting likely locations of prehistoric sites in unsurveyed portions of the project area	"Grand Overall Model": when 80, 90, and 95% of sites are predicted correctly, so are 47, 39, and 29% of "nonsites," respectively (pp. 150-153)	Tables, graphs, and computer-generated maps (at scales from 1:250,000 to 1:24,000), including probability surface maps for site location in tracts ranging in size from ca. 24 to 1 ha (60 to 2.5 acres)	Predictive models generated from map-based survey data were tested independently using site file data, "nonsite" data, and split-sample techniques	5	5	5	3	3

(continued)

TABLE A.1. (Continued)

Summary of general characteristics of the selected predictive model projects

Project Location	*Type of Model*	*Model Objectives and Applications*	*Claimed Level of Accuracy*	*Model Presentation*	*Approach to Verification*	*General Assessment (Scale 1:5)*				
						Data Evaluation	*Falsifiability*	*Consistency*	*Simplicity*	*Generalizability*
Kaibab (*K*) and Cuba (*C*) Study Area Projects (Plog 1983a, 1983b)										
K: North-central Arizona (Coconino County); C: North-central New Mexico (Rio Arriba County)	K and C: Inductive logic in an analytic context	K: To use 1% sample and planning data to make predictions and to use other survey data to determine how accurate predictions would have been, as an experimental predictive model approach; C: To examine feasibility of doing predictive modeling with available data, as an experimental predictive model approach	K: NRO, but low (p. 66); C: data result in "as clear a definition of an approach for finding all sites with less than inventory survey as one can imagine" (p. 78)	K: Tables and plan view maps (ca. 1:140,800); C: Text and plan view maps (ca. 1:42,240)	K: Compare results of two surveys; C: Untested, but could be tested with new survey data	K:3 C:1	K:3 C:5	K:3 C:3	K:5 C:5	K:3 C:1
Fort Benning 2200-Acre Survey Project (Thomas et al. 1983)										
West-central Georgia; Chattahoochee County	Inductive logic in an analytic context	To conduct an evaluation of cultural resources in the 8907 ha project area on the basis of a 10% judgmental cluster sample and to test an earlier model (in Kohler et al. 1980) for management purposes	Discriminant analysis correctly reclassified 96% of cases and distinguished site locations from nonsite locations and correctly reclassified 90% of the sites (pp. 127-129)	Tables, map (1:31,250) illustrating soil types/favorable site locations in survey area and a more detailed map (1:25,000) showing results (point plots) of discriminant analysis test of model against known site and siteless locations	Expected site locations (density by zone) were generated from previous model (Kohler et al. 1980) and new areas were surveyed; then survey results were used in a series of discriminant analyses that refine earlier model and test it against known locations in study area	4	4	4	3	3
Cisco Desert Project (Bradley et al. 1984)										
East-central Utah; Grand County	Inductive logic in an analytic context	To construct predictive model of site and nonsite areas using existing data and to test it with results of survey to identify environmental variables that accurately predict site locations and delineate areas of low site density for management purposes	Soil model yielded predictive accuracy rates from 60 to 85% (pp. 86, 88); discriminant function model yielded accuracy rates less than 80% (p. 95)	Tables and maps illustrating soil type/site densities for the entire Cisco area (1:24,000 and reduced versions for text); also discriminant function coefficients to calculate probabilities for site location in 16 ha (40 acre) areas	Used results of survey to test predictions indicated from existing survey data; used discriminant analysis to distinguish site and siteless areas; compared results of other previous surveys to predictions to generate accuracy rates for soil model	3	4	3	4	3

Route 13 Relief Corridor Project (Custer et al. 1984)										
Central Delaware; Castle and Kent counties	Inductive logic in an analytic context with elements of a deductive, systemic model, especially for mortuary/exchange sites	To identify zones within the study area likely to contain significant prehistoric and/or historical resources for planning and management purposes	Given adequate training sets and ground-truthed LANDSAT data, classification is 87% accurate (p. 80) and site probabilities per site pixel are "reliable and accurate estimates" (p. 130)	Tables, charts, maps (1:24,000) depicting probability zones and sensitivity (regarding overall significance); also generalized probability estimates for site presence per 2.3 ha area	Formal test of the models generated are not conducted, but predictions are compared to other models as an informal test and future surveys can test model	4	3	3	2	3
Montane Hunter-Gatherer Project (Hackenberger 1984)										
Central Idaho; Custer, Idaho, Lemhi, and Valley counties	Deductive logic in a systemic context	To determine whether proportional resource use can be predicted by measures of resource payoffs, by decision strategies or by resource distribution and to determine whether archaeological data can address this problem	Relative effects of resource distributions and changes in them for choices of site locations are greater than effects of different manners of making decisions (p. 141)	Tables, charts, and various maps (ca. 1:833,333); SYMAPs at same scale illustrate density of various food resources, seasonally available calories, predicted proportional winter use of ungulates for each 4828 by 4828 m grid unit (23.3 km²)	Archaeological investigation of site location, resource utilization, and settlement size are discussed with regard to testing decision models; formal test can be made with existing and/or new survey data	4	3	4	3	5
Tar Sands Project (Tipps 1984)										
Southeast Utah; Garfield, Emery, and San Juan counties	Inductive logic in an analytic context	To develop, test, and refine a site locational model based on correlations between environmental characteristics and known site locations for planning purposes	Discriminant model classified 93% of the quadrats with sites (p. 163); LANDSAT model is not immediately practical for management purposes (p. 172)	Tables with predicted site densities per 65 ha (160 acre) quadrat per study tract, and tabular presentation of the discriminant model for site and siteless quadrats	The site density estimates were not tested, but the discriminant model was tested and refined with an additional 5% (area) sample and the final model was based on the 10% combined sample; it was also tested with a split sample	4	4	4	4	3
Central Oregon Project (Wildesen 1984)										
Central Oregon; Crook, Deschutes, Harney, and Lake counties	Inductive logic in an analytic context with a few deductive elements (e.g., use of ethnographic analogy)	To identify and stratify lands likely to contain significant prehistoric sites requiring "affirmative management action" and to identify lands not likely to contain significant sites so the agency can comply with the National Historic Preservation Act by concentrating its management efforts on sites with significant values	7.3% of area is suitable for containing intact sites with significant information; as much as 93% of the study area (ca. 397,853 ha) can be removed from the potential data base (p. 5)	Text, tables, graphs, and maps (scale NRO) of areas with "high probability landforms"; computer printouts environmental data for all areas	Untested, but model is subject to testing with existing environmental and cultural data as well as with new survey data	3	2	2	4	3

TABLE A.2.

Summary of characteristics of data base and predictions from selected predictive model projects

Sample Size			*Environmental Measurements in Model*	*Type of Predicted Resource*	*Prediction/Estimation Procedures*	*Resolution of Prediction*	*Discussion of Procedure (Scale 1-5)*
Sites	*Spatial Units*	*Type of Sample*					
Reese River Ecological Project (Thomas 1973)							
NRO*; focus on artifacts/features	140 500 by 500 m tracts totaling 3500 ha (8645 acres)	10% stratified random sample of grid units	Nominal	Artifact and feature distributions by grid unit	Computer simulation to model expected artifact/feature distribution	500 by 500 m grid units (25 ha), with predictions for material densities and dispersions	4
Elk Ridge Project (DeBloois 1975)							
642	Computer simulation runs for grid units from 150 by 150 m to 1200 by 1200 m	Simple random sample of sites within landform types	Nominal, ordinal	Site density by grid unit	Simple map-based classification of area into landforms	Site density by areas ranging from 150 by 150 m to 1200 by 1200 m	4
Lake Koocanusa Project (Jermann and Aaberg 1976)							
21	23 tracts, each 804 m wide but of varying lengths (area from 8 to 225 ha)	Planned 10% stratified random sample, but rising lake levels limited overall coverage to a 6.1% areal sample	Nominal	Site density by zone	Predicted densities calculated by multiplying known density per unit/area by total area per stratum	Site density per topographic stratum (various sizes)	4
CO_2 Project (Woodward-Clyde Consultants 1978; James et al. 1983)							
NRO*, but 900 cells (122 by 122 m) with sites	1700 122 by 122 m (ca. 1.5 ha) cells totaling ca. 2550 ha (6299 acres)	NRO; 1700 cells represent about 1% of the study area	Nominal, ordinal, ratio	Significant site potential by grid unit	Multiple regression analysis used to find correlation between environmental variables and site location	Site significance measures for 122 by 122 m grid units	4
Continental Shelf Project (Barber and Russell 1979)							
6600	Various	All available data on sites	**IN**ductive: Nominal to ratio; **DE**ductive: nominal	**IN** and **DE**: site size, density, and type by zone	Largely map-based classification of area by environmental zones and judgmental classification of site types, then direct extrapolation to similar inundated areas	**IN**: Generalized probability (low, high) of site-type location by landform in zones of various sizes; **DE**: Generalized prediction of intensity of use in a given zone	4
Fort Benning 4000-Acre Project (Kohler et al. 1980)							
31	One block, ca. 1620 ha	Judgmentally selected block area	Nominal, ordinal interval, ratio	Site density by zone	Variables determining site location were identified using analysis of variance and other bivariate significance tests	Generalized probability or site-likelihood strata (based on site densities in similar surveyed areas) for various areas (ca. 109-1194 ha) are delimited	4
Tombigbee Early Man Project (Moto and Gunn 1980)							
35	Various size locations (termed	37% proportional, stratified random	Nominal, ordinal	Site presence absence by landform	QPs characterized according to five environmental variables;	Probable site locations (QPs) ranging in size from 15 to 90 m	2

	QPs) from 15 to 90 m in diameter; also 500 by 500 m grid units	sample of judgmentally determined strata, coupled with judgmental and kriging samples		or grid unit	QPs with identical variable codes were grouped to form sampling strata; the classification procedure for the kriging model is NRO	in diameter are plotted; in kriging model, grid units (500 by 500 m to 2000 by 2000 m) are plotted	
NPR-Alaska Project (Ebert and Brown 1981)							
478	Various	All site data from all surveyed areas, about 2.5% of total area (an ex post facto sampling design; p. 361)	Nominal	Site density by zone	LANDSAT imagery data used to classify ecologic/cover-type strata; site density predicted on basis of density in surveyed area	Ecologic/cover-types (with known and/or predicted site densities) of various sizes are plotted	4
Seep Ridge Project (Larralde and Chandler 1981)							
40 for sensitivity; 34 for discriminant analysis	274 units, 805 by 201 m (1/2 by 1/8 mile)	10% nonstratified, systematic random sample selected by BLM	Nominal for sensitivity ratio, interval for discriminant analysis	Probability of site presence by grid unit	Simple correlation for sensitivity model; stepwise procedure to reduce variables and then discriminant function analysis to classify site vs nonsite tracts	Generalized site probability zones are plotted on maps for sensitivity model; probability of occurrence of prehistoric sites in 16 ha (40 acre) units can be defined within study area	3
Okanogan Highlands Project (Mierendorf et al. 1981)							
456 sites in project area	NRO; a wide variety of surveys conducted over several decades	All known sites in project area; about 0.3% of uplands and 1.2% of riverine areas have been surveyed	Nominal, ratio	Generalized site density by zone	Delineation of seasonal land-use zones based primarily on vegetation maps and ethnographic data; these zones also incorporate relative site-density estimates based on ethnographic settlement-subsistence data	Various zones (e.g., land use, relative density, potential significance) are plotted on the maps; zones range from 2 by 5 km to 30 by 110 km; land-use zones include summer hunting and gathering, summer fishing, winter residence, etc.	3
Salmon River History Project (Rossillon 1981)							
NRO, but at least 48 homesteads and 4 local markets	520 grid units, each 4828 by 4828 m (ca. 23.3 km²)	No sample taken; entire study area was characterized according to stock-raising potential	Ordinal, ratio	Generalized intensity and type of use by grid unit	Grid units (23.3 km²) were characterized according to the relative proportions of habitat types as a means to determine palatability index values and natural production of grasses; high-use areas equate with high index values and high grass production	Generalized high-intensity use areas for summer and winter camps for sheep- and cattle-raising are described in the text according to elevation and named valleys 10 to 40 km in length	3
Bisti-Star Lake Project (Kemter 1982)							
91 in survey area	Survey of 15 259 ha units; 2 of 194 ha; and 1 of 324 ha; 813 2000 by 2000 m grid units in environmental analysis	A ca. 15% judgmental sample conditioned by ownership, access, budget, etc.	Nominal (LANDSAT)	Site type densities by grid unit	Numerous multiple regressions to reduce number of variables; predictions made by calculating mean frequency for each site type within each environmental zone	Predicted site-type frequencies within 2000 by 2000 m grid units are plotted on maps illustrating environmental zones	2

*NRO: Information is Not Readily Obtained from referenced source.

(continued)

TABLE A.2. (Continued)

Summary of characteristics of data base and predictions from selected predictive model projects

Sample Size: Sites	Sample Size: Spatial Units	Type of Sample	Environmental Measurements in Model	Type of Predicted Resource	Prediction/Estimation Procedure	Resolution of Prediction	Discussion of Procedure (Scale 1:5)
Ozark-St. Francis National Forest Project (Sabo et al. 1982)							
493 sites with 560 components	NRO; a wide variety of surveys provided information in site files	All site files with adequate information were used; no estimate on percent of area surveyed	Nominal, interval	Generalized site densities by zone	Univariate and bivariate statistics are used to reduce number of variables; a Q-mode cluster analysis used to recognize and delineate patterns and define groups of sites with similar biophysical characteristics	Generalized site-likelihood zones are plotted on maps (1:24,000); zones are primarily topographic (e.g., bottom lands, bluff lines) and vary greatly in area	4
Passaic River Basin Project (Hasenstab 1983)							
85 prehistoric sites used to develop model	4306 grid units ca. 61 by 76 m (ca. 0.47 ha)	Systematic stratified random sample (with modifications); overall sample fraction of ca. 6.5%	Nominal, ratio	Probability of site presence by grid unit	A Geographic Information System (GIS) was developed; univariate statistical tests determined which variables were significant; pixel sensitivity measured by values derived from "composite subscores of various significant variables," range of scores divided into strata	Series of computer-generated maps produced by plotting generalized scores (high, medium, low) derived from individual scores for site sensitivity, disturbance, modern land use for each pixel (0.46 ha) along watercourses; results are plotted on 1:24,000 maps	3
Grand Junction Resource Area Project (Kvamme 1983)							
202 new sites	Survey done in 100 tracts, ca. 65 ha	Stratified proportional random sample within three subareas, amounting to a 1.5% survey sample from the overall area	Nominal (biotic zones), interval, ratio	Probability of site presence by grid unit	Models developed through a pattern recognition approach using quantitative multivariate environmental (GIS-based) data and several classification approaches (e.g., logistic regression, a quadratic model, a parallelepiped classifier)	Depending upon particular approach, probability surface maps provide predictions for site and nonsite occurrences for areas ranging from 0.6 to 25 ha	4
Kaibab (*K*) and Cuba (*C*) Study Area Projects (Plog 1983a, 1983b)							
K: 187 C: 142	K: large (486-3428 ha) block areas; C: block area	K: one survey was 1%, another was 100% of direct impact zone in forest; C: NRO, probably 100% of direct impact zone	Nominal	Generalized site density by zone	K: Study area divided into drainage basins and site density (relative and absolute); C: Study area divided into two topographic zones, and site density figures calculated for each zone	K: High site density zones plotted on maps; zones vary from ca. 45 to 2500 ha; C: High site density zones plotted on maps; zones vary from ca. 8 to 340 ha	2
Fort Benning 2200-Acre Survey Project (Thomas et al. 1983)							
37 in surveyed area	Block area of ca. 891 ha	A pre-selected 10% sample (one block	Nominal, ordinal,	Probability of site presence at any	Data from site and siteless locations and from a 225 m	Refinement of previous site-location maps to illustrate	5

		unit) of the 8907 ha maneuver area	interval	location	radius of the locations were used in a series of discriminant analyses to distinguish site from siteless areas and prehistoric from historical sites	favorable zones from 5 to 60 ha each; development of maps that illustrate selected locations of predicted site (historical and prehistoric) and siteless areas	
Cisco Desert Project (Bradley et al. 1984)							
126 in surveyed area	100 transects, ca. 201 by 804 m (16 ha)	A 5% simple random sample	**Discrim:** interval; **Soil:** nominal	**Discrim** and **Soil:** probability of site presence by grid unit	**Discrim:** Elevation, view spread, distance to permanent water, and slope used to distinguish prehistoric site and siteless areas; **Soil:** site densities in soil type zones used to create generalized site-likelihood (i.e., sensitivity) zones	**Discrim:** Calculated coefficients can be used with environmental variables (measurable) to determine probability of site encounter for any 16 ha area; **Soil:** Mean site-density figures for soil units are plotted and designated as sensitivity zones ranging in size from ca. 240 to 8000 ha	4
Route 13 Relief Corridor Project (Custer et al. 1984)							
150 aboriginal sites in general models	For the general models, the area is a block unit; for the LANDSAT ODESSA model, grid units (pixels) are ca. 152 by 152 m or 2.3 ha each	All recorded sites in corridor and all areas in corridor are used in models	Nominal (including LANDSAT classification)	Probability of site size and type by zone	General locational models compared to results of terrain analysis and to known distribution of sites; logistic regression analysis used to correlate environmental zones with site presence and results are generalized to yield site probability zones	Diagrams illustrate relationships among site types and locational settings; generalized probability zones of various sizes and site types are plotted on 1:24,000 maps; tables provide information on site probability significance for historical sites	2
Montane Hunter-Gatherer Project (Hackenberger 1984)							
NRO, selected sites compared to model	520 grid units, each 4828 by 4828 m (ca. 23.3 km²)	The entire study area is considered and sites are selected for comparative purposes on a judgmental basis	Nominal, ordinal, ratio	Generalized probability of site type by grid unit	Grid units were classified according to ecological unit, plant, faunal, and fish resources for various seasons and predictions were made based on ethnographic analogy for proportional resource use	Proportional use of ungulates during winter, seasonal caloric yields for other foods, potential population aggregation as well as other estimates by 23.3 km² area; generalized predictions for the locations of seasonal site types within portions of specific valleys in the study area	5
Tar Sands Project (Tipps 1984)							
155 sites in all, but 54 quadrats with sites were used for model	805 by 805 m or ca. 65 ha	Two 5% simple random samples (with judgmental modifications) of 65 ha quadrats	Interval, ratio	Probability of sites by grid unit	For **density estimates** data are evaluated and population estimates with confidence intervals are made for the three study tracts; the **discriminant model** classified grid units as having either no sites, 1 site, or 2+ sites; the **LANDSAT model** used cluster analysis to classify area and results of survey to provide probabilities for each environmental zone	**Site desity estimates** are made for 65 ha grids and presented in tables and text, as are results of the **discriminant model** and **LANDSAT model**, all results were plotted on maps with sections indicated	5

(continued)

TABLE A.2. (Continued)

Summary of characteristics of data base and predictions from selected predictive model projects

Sample Size: Sites	Sample Size: Spatial Units	Type of Sample	Environmental Measurements in Model	Type of Predicted Resource	Prediction/Estimation Procedures	Resolution of Prediction	Discussion of Procedure (Scale 1:5)
Central Oregon Project (Wildesen 1984)							
244 sites on BLM lands, but 364 sites in overall study area	Block area with private and BLM lands; general unit of analysis is 256 ha	All sites on all BLM land in the study area (about 45% of the land and 67% of the recorded sites)	Nominal, ratio	Generalized probability of significant sites by landform type	Computer sorting by major similarities resulted in 33 landform categories with different soils, vegetation communities, and wildlife habitat; correlations between potential food resources and landforms facilitate the identification of 7 "high probability landforms," or those most likely to contain significant sites	A map (scale is NRO) illustrating areas that meet "high probability landform" criteria was generated by the Automated Cartography group at the BLM state office and was color-coded to illustrate rimrock and mesa edges and ecological habitat types most used by ethnographic populations	4

TABLE A.3.

Summary of environment and environmental variables used in selected predictive models

Physiographic Division/ Current Land Use	*Size of Target Area**	*Variables Used in Model-Building*		*Level of Consideration (Scale 1:5)*		
		Type	*Source*	*Paleoenvironmental Factors*	*Depositional/ Environmental Factors*	*Relevant Ecosystem*
Reese River Ecological Project (Thomas 1973)						
Basin and Range; mostly rangeland	77,733 ha (192,000 acres)	Lifezones	Literature and maps	1	1	4
Elk Ridge Project (DeBloois 1975)						
Colorado Plateau; mostly rangeland, some forests	133,603 ha (330,000 acres)	Landform and management areas	Maps	1	1	2
Lake Koocanusa Project (Jermann and Aaberg 1976)						
Northern Rocky Mountain; reservoir drawdown zone	4806 ha (11,871 acres)	Valley width, side of river, stream proximity	Maps, field observations	2	2	2
CO_2 Project (Woodward-Clyde Consultants 1978; James et al. 1983)						
Colorado Plateau; mostly rangeland, some agricultural land and forests	263,158 ha (650,000 acres)	Soil association, slope, drainage rank	Maps, field observations	1	1	1
Continental Shelf Project (Barber and Russell 1979)						
New England, Coastal Plain, and Continental Shelf; most is ocean floor, but coastline has various uses, from recreation to urban/industrial	315,000,000 ha (788,050,000 acres)	**Inductive:** relief, landform, soil type, drainage, slope, aspect, elevation, type of water, distance to water; **Deductive:** spatial distribution of food resources	**IN:** Literature, maps **DE:** Literature	5	4	5
Fort Benning 4000-Acre Project (Kohler et al. 1980)						
Coastal Plain on boundary with Piedmont Plateau; currently military maneuver area; was range, agricultural lands	74,089 ha (183,000 acres)	Soil type, slope, distance to water	Map, literature	2	3	2
Tombigbee Early Man Project (Mitto and Gunn 1980)						
Coastal Plain; mostly forest and agricultural land	155,465 ha (384,000 acres)	Presence/absence of flowing water, stream confluence, inside of river bend, diverse topography, wetlands at/near location	Maps	4	5	3
NPR-Alaska Project (Ebert and Brown 1981)						
Arctic Slope, Brooks, and MacKenzie Mountains; largely unaltered ("natural") habitat	9,000,000 ha (22,230,000 acres)	Ecologic/cover type defined primarily by plant, soil, and water characteristics	LANDSAT imagery	1	2	2
Seep Ridge Project (Larralde and Chandler 1981)						
Colorado Plateau; rangeland	44,200 ha (109,400 acres)	Sensitivity model used landform and vegetation; discriminant model used presence/absence of dunes, view, distance to wood, relief, shelter, exposure, vantage, and ecotone	Maps for both models	1	1	2

*Spatial population for which predictions are made

(continued)

TABLE A.3. (Continued)

Summary of environment and environmental variables used in selected predictive models

Physiographic Division/ Current Land Use	*Size of Target Area**	*Variables Used in Model-Building*		*Level of Consideration (Scale 1:5)*		
		Type	*Source*	*Paleoenvironmental Factors*	*Depositional/ Environmental Factors*	*Relevant Ecosystem*
Okanogan Highlands Project (Mierendorf et al. 1981)						
Northern Rocky Mountains; rangeland, forests, agriculture, and urban areas	2,166,200 ha (5,350,514 acres)	Primarily vegetation and hydrological features (e.g., falls) are used to define land use zones	Maps and ethnographic literature	3	1	4
Salmon River History Project (Rossillon 1981)						
Northern Rocky Mountains; forests and rangeland	1,212,098 ha (2,995,200 acres)	Percent vegetation coverage, palatability index values for habitat types, and grass production as a measure for hay production/number of animals supported	Maps and literature	1	1	4
Bisti-Star Lake Project (Kemrer 1981)						
Colorado Plateau; mostly rangeland	31,413 ha (77,590 acres)	Environmental classes or zones based primarily on soil and washes (gullies) and combinations thereof	LANDSAT imagery	1	1	2
Ozark-St. Francis National Forest Project (Sabo et al. 1982)						
Ozark Forest: Ozark Plateau and Quachita physiographic provinces; St. Francis Forest: Coastal Plain province; primarily forest but some agricultural lands	Ozark Forest: 452,523 ha (1,118,170 acres); St. Francis Forest: 8477 ha (20,946 acres)	Statistical model: topographic setting, soil class, distance to water, and elevation	Maps and literature (for soils)	5	5	4
Passaic River Basin Project (Hasenstab 1983)						
Atlantic Coastal Plain, Valley and Ridge, and Plateau Piedmont; urban, transportation routes, recreation, agricultural lands, forests, and wetlands	ca. 1619 ha (4000 acres); actually 4951-acre pixel area	Those found to be statistically significant: soil drainage, distance to nearest river, to minor confluence, and to major confluence	Maps	1	2	3
Grand Junction Resource Area Project (Kvamme 1983)						
Colorado Plateau and southern Rocky Mountains; rangeland	438,996 ha (1,084,320 acres)	Biotic zones; vertical distance to permanent water, vantage distance, slope, view, exposure, presence/absence of shelter within 100 m and 250 m	Maps and statistical calculations	2	2	1
Kaibab (*K*) and Cuba (*C*) Study Area Projects (Ploy; 1983a, 1983b)						
K: Colorado Plateau; forest and rangeland; C: Southern Rocky Mountains; forest and rangeland	K: 4858 ha (12,000 acres); C: 3427 ha (8464 acres)	K: Drainage basin, vegetation type; C: General landform	K: NRO**, maps (?); C: NRO, maps (?)	K:1 C:1	K:1 C:1	K:1 C:2
Fort Benning 2200-Acre Survey Project (Thomas et al. 1983)						
Fall Line Hills at boundary of Coastal	8907 ha (22,000 acres)	Discriminant analysis used slope,	Maps	2	2	2

Plain and Piedmont Plateau; currently military maneuver area; was forest, rangeland, and agricultural land	in maneuver area of Ft. Benning	distance to water, soil type, landform, relative elevation, and number of streams; soil types determined by the dominant soil type, number of vegetation communities and dominant type within 225 m of site/nonsite				
Cisco Desert Project (Bradley et al. 1984)						
Colorado Plateau; rangeland	32,387 ha (80,000 acres)	**Discriminant:** distances to primary and secondary water (horizontal and vertical), nearest vantage point, nearest juniper, elevation, slope, and downhill view; **Soil:** six soil units (types)	Maps	1	1	2
Route 13 Relief Corridor Project (Custer et al. 1984)						
Coastal Plain; various wetlands, agricultural lands, transportation routes, urban areas	ca. 72,772 ha (179,200 acres)	Environmental classes: five for wetlands, one for forest, two for agricultural lands, two for "bare soils, dead grasses"	LANDSAT imagery	2	1	2
Montane Hunter-Gatherer Project (Hackenberger 1984)						
Northern Rocky Mountains; forest and rangeland	1,212,098 ha (2,995,200 acres)	Vegetation units; available browse for ungulates; distribution of plant, fish, and animal foods by available categories per unit area on a seasonal basis	Maps and literature	2	1	5
Tar Sands Project (Tipps 1984)						
Colorado Plateau; mostly rangeland	69,635 ha (171,000 acres)	**Density Model** used only the three large survey tracts; **Discriminant Model** used relief, distance to water, percent of piñyon/juniper cover and number of drainages in quadrat	Maps	1	2	2
Central Oregon Project (Wildesen 1984)						
Columbia and Snake River Plateau; mostly rangeland	Overall 946,202 ha (2,337,121 acres), but BLM lands account for only 427,787 ha (1,056,633 acres)	Landform type, soil series and phase, soil depth, associated vegetation community, also fauna	Maintained on computer; also Oregon Automated Ecological Site Information System, ethnographic literature, BLM Ecological Site Inventory	1	2	2

**NRO: Information is Not Readily Obtained from referenced source

TABLE A.4.

Summary of cultural identification and specificity of variables used in selected predictive models

	General Site Types		*Variables Used in Model-Building*		*Consideration of Land-Use Systems (Scale 1:5)*
Culture Area	*Known*	*Predicted*	*Type*	*Source*	
Reese River Ecological Project (Thomas 1973)					
Great Basin	Artifact/features judged to represent hunting, butchering, plant-processing, and habitation sites	Same as "known" site types	Artifact and feature types considered representative of selected activities	Literature, project-specific analyses	4
Elk Ridge Project (DeBloois 1975)					
Oasis (Southwest)	NRO*, emphasis is on sites with architecture and/or ceramics	Emphasis on site presence/absence; some attempt to predict feature types (e.g., slab sites or rooms) on landforms	Architecture type, site size	Literature and site files	1
Lake Koocanusa Project (Jermann and Aaberg 1976)					
Plateau near boundary with Plains	Paleoindian to Late Prehistoric (based on projectile points) to modern	Prehistoric (all types combined)	Site density per unit area	Results of survey	2
CO_2 Project (Woodward-Clyde Consultants 1978; James et al. 1983)					
Oasis (Southwest)	Paleoindian, Archaic, and Anasazi sites (most from AD 450 to 1250)	Combination of site type significance score	Period, size, type, history of multiple occupation, site significance	Site files, related information; significance scale judgmentally determined by team of archaeologists	1
Continental Shelf Project (Barber and Russell 1979)					
East and Coastal Plain	Wide range of site types for all periods, Paleoindian to historical	Camp, rockshelter, farmstead, village, habitation, midden, fish weir	**IN**ductive: location, period, stratification, nature of materials; **DE**ductive: general quantification of energy expenditure vs gain for food resources	**IN**: site forms; **DE**: literature	5
Fort Benning 4000-Acre Project (Kohler et al. 1980)					
East (Southeast)	Perhaps Paleoindian, abundant Archaic and Late Prehistoric sites, also historical sites of various types	Emphasis on historical and prehistoric site presence or absence	Primarily site location; some attention to artifact types	Results of survey	3
Tombigbee Early Man Project (Muto and Gunn 1980)					
East (Southeast)	Focus on known "Early Man" sites (Paleoindian, transitional, and Early Archaic), especially basecamps and small sites	Locations of landforms expected to contain Early Man sites (late Pleistocene/early Holocene)	Empirical model: tool types, site type; Initial predictive model: location of appropriate landform	Empirical model: literature; Initial predictive model: maps	2

NPR-Alaska Project (Ebert and Brown 1981)					
Western Arctic	NRO; scatters, isolates, petroglyphs, rockshelters, various features	Mostly site densities, but also relative percentage of site types and other characteristics by zone	Primarily site location, but also site, artifact, and feature type	Site files	2
Seep Ridge Project (Larralde and Chandler 1981)					
Great Basin	Wide range of time periods and site types	Sensitivity model predicts overall site density by zone; discriminant function model provides probability of any site type occurrence	Site presence/absence	Site files and results of survey	1
Okanogan Highlands Project (Mierendorf et al. 1981)					
Plateau	Wide range of time periods from early postglacial through historical and various descriptive types (e.g., rockpiles, rock art, depression/midden, lithic scatter)	Primarily extractive locations, temporary camps, and winter residences	Location and site type (descriptive)	Site files and reclassification of site types	3
Salmon River History Project (Rossillon 1981)					
Plateau	Military camps, mining sites, trapping sites, ranch sites, logging sites, homesteads	Winter and summer grazing sites for cattle and sheep; remains of permanent and temporary stock-tending sites (AD 1866-1930)	Distance to local market and inferred site type (temporary or permanent)	Maps and literature and site files	4
Bisti-Star Lake Project (Kemrer 1982)					
Oasis (Southwest)	Paleoindian, Archaic, Anasazi, Navajo/historical; large camps, small camps, plant-processing, plant- and lithic-procurement, and hunting sites	Lithic sites; Anasazi sites; pre-1933 Navajo sites; post-1933 Navajo sites; Anglo/Spanish sites; total Navajo sites; total sites	Location and site types (same as predicted types)	Site files and new survey data	2
Ozark-St. Francis National Forest Project (Sabo et al. 1982)					
Prairies and East (Southeast)	All major time periods (Paleoindian, Late Archaic, Mississippian, and historical) and many different site types (e.g., basecamps, villages, homesteads)	A variety of site types indicative of identified "adaptation type models" (e.g., Pleistocene/Holocene hunting and gathering, Late Holocene horticulture) and "Historic Activity Periods" (e.g., Spanish, Localized Industry, Forest Service)	Reliability of site location information and many artifact, size, and locational variables used in deductive model; location used in inductive model	Site files	4
Passaic River Basin Project (Hasenstab 1983)					
East	The 85 prehistoric sites in the area (used to build model) were considered to be "a single population of generalized prehistoric sites"	Site presence/absence by pixel	Site location for prehistoric sites	Site files	1

*NRO: Information is Not Readily Obtained from referenced source.

(continued)

TABLE A.4. (Continued)

Summary of cultural identification and specificity of variables used in selected predictive models

	General Site Types		*Variables Used in Model-Building*		*Consideration of Land-Use Systems (Scale 1:5)*
Culture Area	*Known*	*Predicted*	*Type*	*Source*	
Grand Junction Resource Area Project (Kvamme 1983)					
Great Basin	Hunter-gatherer sites from past 7000 years; mostly "open-air scatters," some rock art and rockshelter sites	Location of any site except rock art and rockshelters and other sites in two of the zones	Site presence/absence per grid unit	Results of survey	1
Kaibab (*K*) and Cuba (*C*) Study Area Projects (Plog 1983a, 1983b)					
K: Oasis (Southwest); C: Oasis (Southwest)	K: Desert Culture, Kayenta Anasazi, Virgin Anasazi, Cohonina peoples; small lithic scatters, masonry structures, rockshelters, rock art; C: Gallina culture only; various kinds of residential structures, towers, reservoirs, checkdams	K: Masonry, artifact scatters, and lithic sites by density zone; C: NRO, probably site density by zone	K: general site location, general site type; C: general site location, limited site type information	K: results of previous surveys; C: results of previous surveys	K:1 C:3
Fort Benning 2200-Acre Survey Project (Thomas et al. 1983)					
East (Southeast)	All time periods represented in region (Paleoindian through historical) and numerous site types	Primarily presence/absence of historical and prehistoric sites	Site location	Site files and results of survey	2
Cisco Desert Project (Bradley et al. 1984)					
Oasis (Southwest)	Various site types representative of Paleoindian, Archaic, Fremont, and Shoshonean culture/time periods	Site presence/absence	Site location	Site files and results of survey	1
Route 13 Relief Corridor Project (Custer et al. 1984)					
East	Various site types representative of Paleoindian, Archaic, Woodland, Contact period, and five chrono/historical periods	Up to seven types (e.g., macroband basecamp, hunting site, quarry site, procurement site) for five prehistoric time periods; similar approach for historical sites	Ten site types for five prehistoric time periods; seven functional types for five historical time periods; also evaluation of data quality	Site files	3
Montane Hunter-Gatherer Project (Hackenberger 1984)					
Plateau	Wide variety of site and artifact types representative of Paleoindian through late Prehistoric and early historical periods	Late Prehistoric and early historical site types and locations according to different decision-making criteria (LaPlace, Savage, and Wald models)	Site location and site types, namely temporary procurement camps (e.g., hunting, fishing, root gathering) and winter settlements	Site files and literature	5

Tar Sands Project (Tipps 1984)					
Great Basin	Various site types representative of Archaic, Anasazi, Fremont, Numic, and Euroamerican culture/time periods	For the site density model, site presence/absence is predicted (excluding Euroamerican sites and isolated finds); other models predict whether quadrats will have 0, 1, or 2+ prehistoric/aboriginal sites	Number of sites in quadrats (0, 1, or 2+)	Results of survey	2
Central Oregon Project (Wildesen 1984)					
Plateau and Great Basin	Caves, shelters, rock art, villages, quarries, rock walls and features, camps, knapping stations, processing sites, and others representing the full range of site types in other parts of the Desert West	Significant sites or those with intact sediments and diagnostic materials	Site location and compilation of flora and fauna used by Native Americans known to have occupied the area	Site files and ethnographic literature	2

TABLE A.5.

Summary of general characteristics of field investigations conducted in conjunction with selected predictive model projects.

Use of New Field Data	*Size of Survey Area*	*Survey Procedures*	*Site Number, Type, Density*	*Discussion of Constraints (Scale 1:5)*
Reese River Ecological Project (Thomas 1971)				
To provide archaeological verification of Steward's model	3500 ha	Survey of grid units (25 ha) at a rate of ca. 6.2 ha per person/day; intensive surface collections; no systematic subsurface examinations	NRO*, but 97% of all materials assigned to Medithermal period, between 2500 BP and historical times	1
Elk Ridge Project (DeBloois 1975)				
To provide data base for a test of random sampling in archaeology	4495 ha	Survey of various kinds of landforms in parallel transects; survey rate/intensity NRO; total collection of ceramics, lithics, and organic remains at small sites, and within transects at large sites; no systematic subsurface examination	Overall site density is 1 site every 7 ha; ca. 77% of sites had ceramics and were assigned to Basketmaker (II - III) or Pueblo (I - IV) periods	1
Lake Koocanusa Project (Jermann and Aaberg 1976)				
To provide data base for predicting site densities in unsurveyed areas	338 ha	Surveyed 804 m wide tracts in parallel transects 30 m wide at a rate of ca. 6.8 ha per person/day; surface collection (grab) of some temporally diagnostic artifacts; no systematic subsurface examination	Overall prehistoric site density of 1 site every 16 ha; prehistoric, historical, and stone feature sites were recorded	4
CO_2 Project (Woodward-Clyde Consultants 1976; James et al. 1983)				
Existing data used to build model, but 140 cells were "visited"	208 ha	NRO; all 140 randomly selected cells (ca. 0.08%) were "visited in the field" as a means of verification of the model	NRO, but the field visit results were such that "the standard error of predicted-to-observed value was identical to the standard error of the model" (James et al. 1983:23)	2
Continental Shelf Project (Barber and Russell 1979)				
Existing data used to build models	NRO	NRO; various surveys with many different methods	NRO, but site density and types are highly variable within large areas	3
Fort Benning 4000-Acre Survey (Kohler et al. 1980)				
To provide data base for model construction	1619 ha	Survey of 30 m wide transects at an overall rate of ca. 8 ha per person/day, including systematic subsurface examination and a total collection of all materials on the surface	Depending upon site likelihood stratum, density ranged from 1 site every 10 ha to 1 every 1194 ha; sites and/or isolated finds represented early Archaic through historical periods	2
Tombigbee Early Man Project (Muto and Gunn 1980)				
To test empirical model and develop initial predictive model	NRO, but 56 localities tested, most between 15 and 90 m in diameter	Surface survey, contour map drawn with 20 m grid, random selection of 15 - 25% of grid unit intersections for subsurface testing with auger/core, judgmental locations and off-site locations also tested	NRO, but 62% of tested localities yielded cultural materials, including projectile points representative of all culture/time periods	3

NPR-Alaska Project (Ebert and Brown 1981)				
Existing site data used to construct model; field checking to ground-truth ecologic/cover-type strata	ca. 225,900 ha surveyed for sites	Survey methods varied considerably in the nine survey blocks; no systematic subsurface testing	NRO, but site densities range from 1 site every 172 ha to 1 site every 5000 ha	2
Seep Ridge Project (Larralde and Chandler 1981)				
Used to build or test models	4437 ha	Survey of 15 m wide transects, overall rates of 14.5-16.7 ha per person/day; stone tools sketched in field, grab sample of diagnostic artifacts	34 sites and 106 isolated finds representing Archaic through historical periods and including campsites and limited activity sites; overall density was 1 site every 130.5 ha	1
Okanogan Highlands Project (Mierendorf et al. 1981)				
Use of existing data to construct model	32,490 ha or about 1.5% of project area	NRO; various procedures for many different surveys	Density figures NRO, but 59.2% of sites on river terraces; site type frequencies are rockpile, 10%; depression (pit, storage, burial, etc.), 7.9%; rock art, 5.3%; lithic scatter, 5%; midden, 47.2%; shell, 1.3%; human bone, 2.4%; midden/human bone 2.4%; depression/midden, 9.3%; midden/rockpile, 1.9%; other, 7%	4
Salmon River History Project (Rossillon 1981)				
Limited use of existing site file data to build model	None related to model	NRO, but results of archival searches and some reconnaissance-level survey work were employed to build model	NRO, but expected material contents of predicted winter and summer livestock-raising sites are described and high-use areas are located generally	1
Bisti-Star Lake Project (Kemrer 1982)				
To test and refine model	4599 ha	Survey of 15-30 m wide transects to achieve "an intensive 100% coverage"; survey rate NRO; point-provenience collection of unique artifacts or materials requiring further laboratory identification; no systematic subsurface examination	92 sites (with 103 components) and 213 isolated finds; overall site density was 1 site every 50 ha; components by time periods as follows: Paleoindian, 0.97%; Archaic, 7.77%; Lithic, 33.01%; Anasazi, 30.10%; and Navajo/historical, 28.15%	1
Ozark-St. Francis National Forest Project (Sabo et al. 1982)				
Existing data used to construct or test models	NRO	Various procedures for many different surveys and site information reported by amateurs	NRO for individual surveys, but of the 316 components with reasonably reliable information, 79% are prehistoric, 11% are historical, 9% are unknown; of the 251 prehistoric components, 51% are unknown, 3% are Paleoindian, 27% are Archaic, 10% are Woodland, and 9% are Mississippian	5
Passaic River Basin Project (Hasenstab 1983)				
To test the model derived using GIS data and univariate analysis and to ground-truth validity of GIS environmental information	130 ha	Survey of 0.46 ha (1.15 acre) pixels in parallel transects 15 m wide; systematic subsurface examination (shovel tests) in the pixel quadrant "most likely" to have buried materials; survey ratio NRO; type of collection NRO, but surface and subsurface collections were made	Sample survey to test model yielded 28 (64%) historical sites and 16 (36%) prehistoric sites; overall site density was 1 site every 3 ha; density of prehistoric sites was 1 every 8.1 ha	2

*NRO: Information is Not Readily Obtained from referenced source.

(continued)

TABLE A.5. (Continued)

Summary of general characteristics of field investigations conducted in conjunction with selected predictive model projects.

Use of New Field Data	*Size of Survey Area*	*Survey Procedures*	*Site Number, Type, Density*	*Discussion of Constraints (Scale 1:5)*
Grand Junction Resource Area Project (Kvamme 1983)				
To build predictive model	6478 ha	Survey in parallel transects 20-25 m wide at rates of approximately 16-22 ha per person/day; no collection; artifacts recorded on forms, some were sketched; no systematic subsurface examination	Total of 197 "open-air sites" and 5 rockshelters from survey area; overall site density is 1 site every 32 ha	1
Kaibab (*K*) and Cuba (*C*) Study Area Projects (Plog 1983a, 1983b)				
K: Existing data from one survey used to build model; data from other surveys used to test model; C: Existing data used to assess potential of predictive modeling	K:NRO, perhaps 5006 ha; C:3427 ha	K: NRO, probably different methods for different surveys; C: NRO, perhaps different methods for different surveys	K:Surveys yielded 187 sites with density ranging from total absence to 1 every 43 ha, depending on drainage/vegetation zone; C:Surveys yielded 142 sites with overall density of 1 site every 24 ha and a ridge-top density of 1 site every 5.7 ha, all representative of Gallina phase	K:1 C:1
Fort Benning 2200-Acre Survey Project (Thomas et al. 1983)				
To test and refine existing model and to develop a discriminant analysis model	891 ha	Survey in parallel transects 30 m wide; systematic clearing of forest litter and subsurface examination; collections from shovel tests; systematic surface collection; completion of forms for nonsite areas; survey rate NRO	Total of 37 sites and 32 isolated finds, including 20 prehistoric sites, 15 historical sites, and 2 prehistoric/historical sites; overall site density is 1 site every 24 ha; prehistoric site density is 1 site every 40.5 ha	3
Cisco Desert Project (Bradley et al. 1984)				
To build a predictive model	1619 ha	Survey in parallel transects 15 m wide; collected samples (grab?) of diagnostic artifacts and "obsidian and chert source material"; no systematic subsurface examination; survey rate NRO	Total of 111 prehistoric sites; diagnostic artifacts representative of all culture/time periods; 44.1% ceramic and/or lithic scatters, 48.7% campsites, 5.4% rockshelters, 1.8% quarry sites, overall prehistoric site density is 1 site every 14.6 ha	1
Route 13 Relief Corridor Project (Custer et al. 1984)				
Existing data used to construct model and compare results with those of other models	None related to model	Various surveys with various methods resulting in various qualities of data	Of 150 known prehistoric sites: 9.3% Archaic, 48.7% Woodland I, 20.7% Woodland II, 21.3% unknown; for generalized functional types: 7.3% macroband sites, 26% microband sites, 16% procurement sites, and 50.7% unknown; overall known prehistoric site density is 1 site every 485 ha	3
Montane Hunter-Gatherer Project (Hackenberger 1984)				
Existing data used to estimate the potential of archaeological data to test the economic decision model	None related to model	Various surveys with various methods	NRO; mentions upland and river valley/canyon sites as well as rockshelter and cave sites and their potential to yield data for testing models	3

Tar Sands Project (Tipps 1984)				
To develop, test, and refine models	7004 ha	Survey in parallel transects 15 m wide; limited systematic subsurface examinations ("probed"); site forms, sketch maps; tools drawn; collection plotting of potentially diagnostic artifacts; survey rate of about 13 ha per person/day	Total of 167 components (155 sites): 5.4% Euroamerican, 1.8% Numic, 10.2% Anasazi, 3.6% Fremont, 17.4% Archaic, and 61.7% unknown; of the 158 prehistoric components: 15.2% limited activity sites, 47.5% field camps, 31.0% basecamps, 6.3% habitation camps; overall site density is 1 site every 45 ha	1
Central Oregon Project (Wildesen 1984)				
Existing data used to develop model	None related to model	NRO, but probably numerous surveys using different methods	Total of 364 sites in study area; culture/time period information and functional/typological data NRO; site density for 259 ha (640 acre) areas ranging from 1 site every 259 ha to 1 site every 15 ha	1

Basin and range topography, arid sagebrush flats and piñon-juniper woodlands are characteristic of the 77,730 ha project study area in the upper Reese River Valley of central Nevada. Steward's (1938) ethnographic model for the Reese River Shoshone subsistence patterns was tested using archaeological data. Ethnographically derived seasonality, resource use, activity, assemblage, and settlement information was quantified, and the resulting data were used in a computer-based simulation model. Initially the model was used to predict the nature of food procurement and maintenance activities in different environmental zones. Ultimately, the model also predicted artifact and feature distributions and densities in four "lifezones."

The simulation-generated predictions were tested using new survey data. The project area was stratified on the basis of vegetation communities or microenvironmental zones that were exploited differentially by the Shoshone. A 500 by 500 m grid was superimposed on the study area, and a 10 percent sample was selected from each stratum. The resulting 140 grid units (25 ha each) were surveyed, and the locations of individual artifacts and features were plotted on maps; these artifacts and features (rather than clusters defined as "sites") served as the unit of information. Artifact and feature distributions and densities derived from the survey data (see Thomas 1975) were compared with distributions and densities predicted by the simulation model. Finally, statistical significance tests (e.g., chi-square and Mann-Whitney U) were used to examine the relationship between expected and observed values. Steward's model was supported by the survey data in that 75 percent of the predicted frequencies were verified by the archaeological remains.

Given the stated objectives, this project was a successful predictive modeling effort, and the results contributed to existing knowledge because the nature and distributions of cultural resources were defined and partially explained. The ptoject also employed an innovative survey strategy—the nonsite approach—wherein the distributions of artifacts and features across the landscape rather than concentrations of materials (sites) are monitored. That approach circumvents some of the adverse effects that can result from using observed densities of artifacts to distinguish arbitrarily between isolated finds and sites in an attempt to understand past human behavior. The model is subject to criticism, however, for its heavy reliance on the ethnographic tecord. That approach can only be justified insofar as it can be demonstrated that relevant aspects of the ethnographically documented land-use systems are consistent with human behavior in the area during the last 4500 years.

Elk Ridge Project *The Elk Ridge Archaeological Project: A Test of Random Sampling in Archaeological Surveying.* Evan I. DeBloois. Cultural Resources Report No. 2. USDA Forest Service, Intermountain Region. 1975

The Elk Ridge Project was sponsored, in part, by the Forest Service's Intermountain Regional Office as a feasibility study for determining the validity and reliability of random sampling designs in archaeological survey. It was carried out initially by Forest Service personnel and subsequently by individuals representing Brigham Young University. Its objective was to determine whether a predictive

sampling strategy could be designed and implemented as an interim step in the total inventory of a project area. The author was interested in investigating "the reliability of sampling in predicting attributes of the larger population," and specifically in addressing the question of "how many sites can be expected in such-and-such an area?" (DeBloois 1975:4, 126). This project clearly has management-oriented objectives, but it also had research objectives as a study of the utility of sampling in cultural resource management. An early version of the study was a dissertation project at the University of Washington. The information summarized here is from DeBloois (1975).

The study focused on a 133,603 ha area in southeastern Utah comprising ponderosa pine, piñon-juniper, oak-serviceberry, and cottonwood vegetation zones. Some 640 sites were recorded during the survey of a 4495 ha sample of the project area. Almost all of the sites were either habitation or special-use sites assigned to the Basketmaker/Pueblo sequence. Environmental (e.g., vegetation, soil, and landform types) and cultural (e.g., site size, type, and cultural affiliation) data for the study area were coded using a Topcart digitizer. Various random samples of different proportions and quadrat sizes were drawn from the area surveyed and used to calculate the total number of sites. Resulting estimates were compared with the actual data base and assessed using the chi-square test. Assessments were made in an attempt to measure the accuracy of different sampling techniques and sizes. Simple random sampling was found to be a reliable predictor of total population but not necessarily of the distribution of certain site characteristics. When survey of an "unknown area" was simulated and a random sampling scheme was applied, units between 600 and 800 m sq (ca. 36–64 quadrats) were found to be most effective. Because of the "dangers of improper stratification of an unknown population" it was concluded that simple random sampling might be a "better choice" for initial surveys (DeBloois 1975:126).

The Elk Ridge Project was one of the earliest attempts to apply the concepts of sampling and predictive locational modeling to federally mandated cultural resource management. Given that this project served as a prototype, the relatively simple (largely univariate) statistical approaches used cannot be expected to compare favorably to more recent modeling efforts, with their rigorous use of complex multivariate statistics. As is the case with many predictive models generated using data bases where known sites represent only the last few thousand years of prehistory, one is left wondering about the locations of sites representing the preceding 10,000 years of prehistory in the Elk Ridge area.

Lake Koocanusa Project *Archaeological Reconnaissance in the Libby Reservoir–Lake Koocanusa Area, Northwestern Montana.* Jerry V. Jermann and Stephen Aaberg. Department of Sociology, Montana State University. 1976

The Seattle District Corps of Engineers sponsored the Lake Koocanusa reconnaissance project, which was carried out by personnel representing Montana State University. It was funded because Corps personnel discovered a number of previously unrecorded sites in the denuded drawdown zone of Lake Koocanusa in

northwestern Montana. The primary objective of the project was to obtain estimates for the total number, nature, and distribution of sites that might be present in the reservoir drawdown zone (Jermann and Aaberg 1976).

For sampling purposes the 4806 ha, 80 km long study area in the Kootenai River Valley was subdivided on the basis of topography. A series of survey tracts 800 m in width, and of various lengths, were selected randomly from each topographic stratum. The tracts represented between 3.6 and 8.2 percent of the eight subdivisions and totaled about 339 ha or approximately 6 percent of the project area. Twenty-one prehistoric sites, identified as spanning the early Middle Prehistoric (i.e., Archaic) to the Late Prehistoric periods, were documented. Euroamerican sites were also recorded. Site density figures were calculated for the surveyed portions of the various topographic subdivisions, and these figures were multiplied by the total area in each stratum to estimate the total number of sites in the project area. It was estimated that as many as 400 sites might be present in the drawdown zone.

This project is an early example of what might be termed the "direct extrapolation of site density" approach to predictive modeling, or what Kohler and Parker (1986) call *projection.* Although very simple in its approach, this application can be considered successful because with this projection of high site densities the Corps was able to justify funding intensive surveys. In an area where the vast majority of known sites represent only the last few thousand years of occupation and are situated in valley bottoms, the detection and prediction of older sites located well above the valley bottom is recognizably a contribution of information important to our understanding of local prehistory.

Wasson Field–Denver Unit CO_2 Project Predicting Site Significance: A Management Application of High-Resolution Modeling. S. E. James, R. Knudson, A. Kane, and D. Breternitz. Paper presented at the 48th Annual Meeting of the Society for American Archaeology. 1983; "Appendix E," in *Well Field Development Plan for the Wasson Field–Denver Unit CO_2 Project Environmental Impact Report.* Woodward-Clyde Consultants. 1978

The Wasson Field–Denver Unit CO_2 predictive modeling project was funded by a private oil company as part of its effort to develop an environmental impact statement (EIS) for a carbon dioxide well-field project in southwestern Colorado. The cultural resources portion of the EIS was necessary in part because the Bureau of Land Management required a right-of-way permit. Personnel representing Woodward-Clyde Consultants were responsible for preparing a planning study that would improve well-field layout by minimizing impacts to significant archaeological sites. Information summarized here is taken from two draft documents (James et al. 1983; Woodward-Clyde Consultants 1978).

The 263,158 ha project area comprises plateaus and canyons, agricultural land, rangeland, and forests. Environmental and cultural data were entered, compiled, analyzed, and displayed using a geographic information system. Map-based information for land use and soil association, prehistoric farming areas, topography,

roads, archaeological sites (including data on period of occupation, size, type, and condition), biological communities, and geologic materials was coded and digitized for 175,000 cells, each representing ca. 1.5 ha. Site significance was identified as the dependent variable and defined in part on the basis of age, type, size, and number of components for hundreds of known Basketmaker, Puebloan Anasazi, and post-Anasazi sites. A fundamental aspect of this definition of significance was described as the "subjective attitudes of professional archaeologists" (Woodward-Clyde Consultants 1978:E-4). The investigators developed a seven-point scale that they believed conformed to "prevailing opinions of the professional archaeological community" (James et al. 1983:17). Ultimately, three independent environmental variables—soil, drainage rank, and slope—were used in a step-wise multiple regression, with the computed site significance values serving as the dependent variable. Sets of surveyed cells without sites were also included in the analysis. The analysis yielded significance values for each cell, and the scaled values were then color coded and plotted on 1:24,000 scale maps. A total of 140 randomly selected cells were field inspected as a means of verifying the model. The model was supported to the extent that the "standard error of predicted-to-observed value was identical to the standard error of the model" (James et al. 1983:23).

This project serves as an example of a management-oriented model designed to minimize uncertainties and delays in the permitting process. It is innovative in its attempt to define significance by relying on the expertise of individuals knowledgeable about the most abundant kinds of sites in the project area, namely those considered to have been occupied by Anasazi groups between AD 450 and 1250. What might be of concern, at least to archaeologists who specialize in hunter-gatherer studies, is that Archaic sites and Basketmaker II sites were assigned the same code for period of occupation. Furthermore, there is no other provision for isolating site types that may represent some of the limited "evidence of seasonal and sporadic presence of peoples from the Paleo-Indian and Archaic periods (10,000 BC–AD 450)" (Woodward-Clyde Consultants 1978:E-8).

Continental Shelf Project *Archaeology and Paleontology.* Summary and Analysis of Cultural Resource Information on the Continental Shelf from the Bay of Fundy to Cape Hatteras, Final Report, Vol. II. Russell Barber and Michael E. Roberts. Institute for Conservation Archaeology, Peabody Museum, Harvard University. 1979

Personnel representing the Institute for Conservation Archaeology at the Peabody Museum conducted the Continental Shelf Project for the Bureau of Land Management. The project was designed primarily to provide the BLM with information about known or expected prehistoric sites and historically important shipwrecks and to generate predictions about where specific types of sites will be found. Information presented here focuses on the prehistoric sites portion of the study by Barber and Roberts (1979).

Continental shelf, coastal, and nearby low-elevation terrestrial areas between Maine and North Carolina constitute this project's 32,388,664 ha study area.

Inductive site-locational models were generated from the site records for dryland areas similar to areas on the continental shelf. Deductive models were generated for the intensity of settlement in a given zone by relying on knowledge and assumptions about human foraging behavior and relevant paleoenvironmental conditions. General and specific predictions derived from both kinds of models were combined to form a final model. This model was based on the results of a generalized assessment of goodness of fit between predictions. The final model presented generalized predictions (i.e., high, medium, medium-low, very low likelihood) for site size, site density, and site type for each of six 3000-year periods and four environmental zones. The time periods correlate roughly with different sea levels and the resulting changes in the positions of the coastline. These changes differentially affected the distribution and nature of the estuarine, inland valley, upland, and coastal environmental zones in each of the three identified subareas—Maine, southern New England, and Mid-Atlantic. The model's end product is presented on a series of 1:250,000 scale maps that illustrate 122 archaeology zones, each of which is characterized by time period for predicted site types as well as generalized site frequencies and site sizes.

The authors' claim that the project represents an advance in the state of the art of predictive modeling for the nature and distribution of prehistoric sites is justifiable, although the spatial resolution of prediction is low. By combining existing site-file data for some 6600 sites with the theory of optimal foraging strategy and including information derived from environmental reconstruction, paleoclimatology, and other disciplines, the investigators were able to predict and partially explain the distribution of cultural resources. They have provided planners with information on the predicted nature and distribution of sites in a vast area within subdivisions as small as 50 km^2. At the same time, the approach is reasonably compatible with contemporary theories, and it affords the opportunity to discover previously undocumented kinds of cultural resources (for additional discussion, see Chapter 2).

Fort Benning 4000-Acre Project *An Archaeological Survey of Selected Areas of the Fort Benning Military Reservation, Alabama and Georgia.* T. A. Kohler, T. P. DesJeans, C. Feiss, and D. E. Thompson. Remote Sensing Analysts. 1980

Remote Sensing Analysts, a private firm based in Tucker, Georgia, conducted the Fort Benning project for the U.S. Army. The scope of work and contract were developed and administered by the Heritage, Conservation and Recreation Service, Interagency Archeological Services, Atlanta. That agency was responsible for selecting the survey tract and specifying the development of a predictive model to serve as an interim management tool. The source of information for the site predictive model summarized here is Kohler et al. (1980).

Fort Benning, located in the Fall Line Hills portion of east-central Alabama and west-central Georgia, encompasses coniferous and mixed forests, scrub oak and brush, and swamp vegetation zones. A judgmentally selected 1619 ha area was surveyed, and 31 sites were identified. Of these sites, 10 had historical nonaboriginal

components and six had historical aboriginal components; three had Mississippian components; three had Late and Middle Woodland components; eight had Early Woodland and late Archaic components; and three had middle and early Archaic components. Analysis of variance, goodness of fit, and *t*-tests were used to identify soil type, slope, and distance to water as variables that were correlated with site location. Several soil types, slopes of less than 10 percent, and areas between 75 and 225 m from water were identified as favorable site locations throughout the project area. These locations were plotted on 1:25,000 scale maps. The locations of predicted site-likelihood strata (maximum, intermediate, and least likely to contain sites) were defined on the basis of the number of intersecting favorable states and were plotted on other maps. For example, areas with Cahaba sandy loam soil on slopes of less than 10 percent and between 75 and 225 m from water were identified as part of the maximum likelihood stratum, whereas areas with similar soils but on steeper slopes and lying more than 225 m from a creek were defined as part of the zone least likely to contain sites. The model also included site-density estimates for the unsurveyed strata, and it included probability estimates for encountering a site within any given randomly selected area.

The project can be considered successful in that a readily testable model was generated to predict the probability of encountering a site anywhere in the project area. It is noteworthy that this project represents an early and comparatively rigorous attempt to use statistical approaches along with new field data generated as a result of a systematic surface and subsurface survey. As in many of the inductive or correlative models, most of the site-type information, which can be informative about the potential of a site to yield important information, is lost when the various kinds of prehistoric sites are merged to generate a site/siteless dichotomy for predictive purposes. Although the concept of site significance is not directly dealt with in the model, there is an implication that maximum site likelihood zones have the highest probability of containing significant cultural resources, especially larger residential sites. Other kinds of sites that may have potential to yield important information are likely to be encountered in the zones that are least likely to contain sites, and by implication these sites are not as likely to be discovered. For example, some types of vegetal procurement sites might be expected to occur on stony soils far from water.

Tombigbee Early Man Project *A Study of Late Quaternary Environments and Early Man Along the Tombigbee River, Alabama and Mississippi, Phase I.* Guy R. Muto and Joel Gunn. Benham-Blair and Affiliates. 1980

The environmental division of Benham-Blair and Affiliates, an architecture and engineering firm, designed and implemented the Tombigbee Early Man Project. It was funded and partially administered by the Corps of Engineers, but the scope of work and project review were primarily the responsibility of the Heritage, Conservation and Recreation Service, Washington, D.C., and Interagency Archeological Services, Atlanta. The draft report (Muto and Gunn 1980) was the source of information summarized here.

The 77,733 ha project area lies within the Tombigbee River Valley of eastern Mississippi and western Alabama. Forests and agricultural crops cover the alluvial terraces, and swamp vegetation occupies the extensive flood basin. The project's major goal was to develop a model that would predict the locations of Paleoindian and early Archaic sites. Since most of the Early Man sites were expected to be deeply buried in late Pleistocene and/or early Holocene deposits, an important aspect of the project was the prediction of locations of landforms old enough to contain early sites. Toward that end, a generalized "empirical" site-location model was developed based on the known distribution and nature of Early Man sites as well as on inferred Pleistocene and early Holocene environmental conditions. Using the resulting locational criteria (e.g., inside of river bend, near confluence, near wetlands) as predictive variables, the researchers visually scanned topographic maps for likely site locations; 620 such locations, termed *Quaternary projections*, were identified.

A second inductive model was developed using a computer-based "prospecting technique" known as kriging (Muto and Gunn 1980:4–18; see also Chapter 3). The kriging model predicted the location of landforms or areas likely to contain early sites. Toward that end, during the kriging operation the computer searches its data banks for grid units encompassing landforms with environmental characteristics like the landforms known to contain sites. The program provides probability estimates for the likelihood that a given grid unit may contain the appropriate landform. Those grid units predicted to contain sites on the basis of the kriging model were termed *machine projections*.

Both models were tested by on-the-ground examinations of a sample of the Quaternary projection locations and machine projection units. Techniques designed to detect buried sites in lowland and swampy environments were used to determine site presence and absence at the sampled locations. These techniques included the use of soil augers capable of penetrating and recovering several meters of clay-rich sediments, which were examined for the presence of artifacts and chemically tested to detect paleosols or those deposits with the potential of containing cultural materials. A total of 56 Quaternary and machine projections were selected and tested for the presence of cultural materials. Of those, 34 locations were selected using a proportional stratified random sampling scheme. Strata were defined as combinations of locational criteria. For example, one stratum included only locations near stream confluences and wetlands while another stratum included locations with the same criteria plus being on the inside of a river bend. The other 22 locations were selected on a judgmental basis for on-the-ground testing because they exhibited unique environmental characteristics or because they filled spatial gaps in the random sample.

The overall approach achieved some success in that slightly more than half of the randomly selected Quaternary projections yielded cultural materials. This success rate is actually quite high given that few of the sites would have been detected by examination of surface or near-surface deposits. An important contribution of this project to predictive modeling is its extensive use of paleoenviron-

mental data and its attention to depositional processes that alter the shape of the landscape and bury archaeological sites. Although more than 600 locations were identified and almost 10 percent were tested, the resulting data were not employed to calculate probability estimates for detecting a site at any given location or within any stratum. The greatest problems encountered in summarizing this project from the information presented in the draft report were that the detailed discussions of some of the project methods were difficult to understand, and the relationships between these methods and the results of the project were not always clear. In addition, the lack of a detailed discussion about the derivation and use of the kriging model was disappointing. These organizational problems may be a result of the draft status of the project report and could be resolved as part of the editorial process.

NPR-Alaska Project "Remote Sensing in the NPR-A Cultural Resources Assessment." James I. Ebert and Galen N. Brown. In *Anthropological and Historic Preservation Cooperative Park Studies Unit Occasional Paper* No. 25, pp. 349–419. University of Alaska. 1981

The National Petroleum Reserve–Alaska Project was sponsored by the National Park Service, Washington, D.C., and implemented largely by personnel representing the Anthropology and Historic Preservation Cooperative Park Studies Unit of the University of Alaska. Remote sensing components of the project were carried out by personnel from the National Park Service's Remote Sensing Division, Southwest Cultural Resources Center, Albuquerque. The objective was to use a remote sensing approach to correlate environmental settings with known site locations in an effort to increase the accuracy and cost efficiency of the cultural resource assessment of the 9 million ha project area. The report on the remote sensing aspects (Ebert and Brown 1981) provided the information summarized here (see also Chapter 9).

Moist tundra, wet tundra, alpine tundra, high brush, and waterways constitute the basic ecosystems in north-central Alaska, where this project was located. The area includes portions of the Brooks Range, Arctic Foothills, and Arctic Coastal Plain physiographic provinces.

Landsat and high-altitude color infrared imagery data were used to define six ecologic/cover types and six transitional types. These 12 strata and all previously recorded archaeological sites were plotted on 1:250,000 scale maps. Area measurements were made for each stratum, and the amount of land that had been surveyed within the various strata was calculated. Next, cultural, landform, and ecologic/cover-type data were recorded and correlated as a means of characterizing the occurrence or nonoccurrence of site-specific cultural and landform data in each stratum. For predictive purposes, the observed site density in the surveyed portions of each stratum could be multiplied by the area of any unsurveyed parcel (in the same stratum) to determine a site-frequency estimate for that part of the project area. Using similar extrapolation techniques, the project personnel gener-

ated sample data to estimate the relative frequencies of site types and the expected content in unsurveyed areas.

The model is appealing largely because of its simplicity and probable cost-effectiveness as a first-step approximation of the nature and distribution of cultural resources in a vast area. It provides an idea of the number, content, size, and other characteristics of sites that might be expected in an unsurveyed area—information critical to realistic estimates of the time and money required to conduct on-the-ground surveys. As recognized by the authors of the report, however, the predictions are conditioned by the quality of the data, which varied from survey to survey. Furthermore, the approach is unlikely to be particularly useful in predicting the presence of theoretically expected but as yet undocumented kinds of sites because it relies entirely on information about previously discovered site types. It is also apparent that the model would not be of great use in gaining information about past environmental or cultural conditions or about why cultural materials are distributed across the landscape in particular patterns—limitations also recognized by the authors.

Seep Ridge Project *Archaeological Inventory in the Seep Ridge Cultural Study Tract.* Signa L. Larralde and Susan M. Chandler. Nickens and Associates. 1981

The Bureau of Land Management funded the Seep Ridge Project, which was carried out by personnel from Nickens and Associates, a private archaeological consulting firm in Montrose, Colorado. Objectives of the part of the project with which this summary is concerned were *(a)* to derive a formula that would determine the probability of site occurrence at any point in the project area, and *(b)* to delineate for management purposes areas suspected to contain an extremely low density of sites. The authors noted the possibility that "project-by-project cultural resources clearances may not be necessary" in some portions of these extremely low density areas (Larralde and Chandler 1981:1).

Semiarid canyons, ridges, eroded buttes, and dune fields are characteristic of the 44,292 ha project area, as are juniper, sagebrush, grasslands, and some desert riparian vegetation. The BLM used a 10 percent nonstratified, systematic random sampling scheme to preselect 274 16 ha tracts for survey. Within that area, 40 sites and 106 isolated finds were recorded; these remains represent all major occupations of the area, from Paleoindian to Euroamerican. A discriminant function analysis was used to compare the relationships between site and nonsite locations on the basis of environmental attributes—presence/absence of sand dunes, viewspread, distance to vantage points, distance to juniper forest, and a measure comparing site or nonsite vegetation with surrounding vegetation. High, medium, and low sensitivity zones were delimited, primarily on the basis of positive correlation between high density and increasing proximity to juniper trees and sand dunes.

The discriminant equation used in this project is described as a "powerful management tool" because it requires data from only six variables and because values for these variables can be measured for any point on a USGS topographic

map. When values for these variables are "plugged into" the formula, the result is a probability estimate for site occurrence. The authors suggest that "if the probability of site presence is low, archaeological clearance could be granted without the necessity of a field check. If, however, the probability of site presence is near the 50 percent range, a field inventory would be in order" (Larralde and Chandler 1981:136). The authors stress that their results are intended as an example of the power of the technique and that this particular equation should not be used in the planning unit "until the function is strengthened by the inclusion of more data" (Larralde and Chandler 1981:136).

Given the 97.1 percent "accuracy rate" claimed for one version of the discriminant analysis, which classified only one of the 34 sites as a nonsite (Larralde and Chandler 1981:133), the modeling project appears to have successfully achieved its stated objective. There are, however, several potential problems with the model, two of which are noted here. The first problem is that the preselected one-half by one-eighth mile transects do not represent a random sample of the landscape in the project area because the central portion of each quarter-section had no chance of being selected (Berry 1984). The linear transects were "situated in quarter sections so that cadastral monuments could be used to maximize location control. . . . Each sample unit was systematically placed in its quarter section to extend from section corner to quarter corner" (Larralde and Chandler 1981:4).

The second potential problem concerns the equation of zones of low site density with nonsignificance, that is, with areas that merit no further attention. Because of this equation there is no opportunity to determine whether scientifically important cultural resources are present in the low-density zone. It is clearly possible that low site-density zones were occupied at some point in the past when environmental conditions were different and human population was low. Given the procedures summarized above, there would be little chance that old and rare sites would be discovered.

Okanogan Highlands Project *A Cultural Resources Predictive Land Use Model for the Okanogan Highlands.* R. R. Mierendorf, T. K. Eller, D. Carlevato, and P. A. McLeod. Cultural Resources Group Report No. 100-2. Eastern Washington University. 1981

The Bonneville Power Administration, Portland, Oregon, funded the Okanogan Highlands overview/predictive modeling project for an area in north-central Washington. The project was designed to evaluate possible disturbances to archaeological sites along proposed transmission lines, and it was implemented by the Bonneville Cultural Resources Group, Eastern Washington University. This summary, based on Mierendorf et al. (1981), focuses on the prehistoric and ethnographic aspects of the land-use model.

Low, forest-covered mountains and steep-walled valleys with steppe vegetation are characteristic of the 2,166,200 ha study area. Existing site-file data were available for 459 sites representing all major periods of occupation (Paleoindian

through early historical). A predictive model of prehistoric land use was developed based on the seasonal and spatial distribution of resources and on ethnographically documented Native American settlement patterns and subsistence practices. In the report the model is presented as a series of maps that delimit seasonal activity areas and expected site densities (high to low); the latter are based on known site densities in similar areas. Examples of zones delimited on maps include "winter residence areas with moderate site density" and "summer hunting and garhering areas with the lowest site density." A sensitivity analysis was conducted to evaluate construction impacts; it used the predictive model to assess potential site significance according to numeric values assigned for regional research significance, site density, and known impacts to cultural resources. Six sensitivity zones, which correspond to generalized geographic strata, were plotted on maps.

The model provides considerable information about the general location of different kinds of sites but almost no information about the probability of encountering a site at any specific location. Even so, it permitted initial estimation of possible disturbance to sites that would result from construction of a powerline across the project area. The authors note that important sites could occur in the one "low site density/low sensitivity zone" and in two of the low density/moderate sensitivity zones that they have defined, but they consider the probability of encountering such a site along a powerline to be low. They expect that "future sutveys [in the low density/low sensitivity zone] will locate sites that are regionally important" (Mierendorf et al. 1981:117).

Reliance on the ethnographic record to predict prehistoric land-use patterns considerably reduces the generalizing power of the model. The authors recognize one aspect of this problem when they suggest that changing resource distributions might have caused changes in the location of activities. What they do not seem to recognize is the probability that at times in the past, especially when human population densities were much lower than those of the ethnographic present, it is likely that different land-use systems operated. For example, one would expect different distributional patterns for different site types depending on whether people spend the winter near stored foods or depend on frequent moves among areas where food resources are available. In the latter case, winter village sites might not be located in the riverine zone, and fishing sites might not be nearly as common as they were during the ethnographic period. If the prehistoric winter pattern was one of frequent residential moves, a number of small, short-duration residential sites might be located at some distance from the river. Overly heavy reliance on the ethnographic record in developing predictive models could result in cultural resources representative of a very different land-use system remaining undetected.

Salmon River History Project *An Overview History in the Drainage Basin of the Middle Fork of the Salmon River.* Mary P. Rossillon. Cultural Resources Report No. 6. USDA Forest Service, Intermountain Region. 1981

Historical research conducted for the Salmon River project was done by personnel representing Washington State University and the University of Idaho.

The study was sponsored as a joint venture involving these universities, the Idaho State Historical Society, the Forest Service, and the Idaho State Historic Preservation Office (Knudson et al. 1982). Early in the project the researchers recognized that little information was available about the stockmen's culture in the central Idaho area. As one means of acquiring that information, a model was developed to predict the locations of nineteenth-century stockraising-associated sites. Information summarized here was taken from Rossillon (1981).

Mountains and upland valleys are characteristic of the 320,000 ha study area. Coniferous forest, some of which is relatively open, is the dominant vegetation zone, followed by grasslands and meadows. The entire study area was subdivided into 3 by 3 km grid units, and each unit was characterized according to its distance from a local market, the palatability of summer and winter range for cattle and sheep (a calculation based on the percentage of readily accessible fodder), and expected hay production (based on the number of cattle and sheep that could be supported). High-use areas—those with the greatest potential for grazing and hay production and those with the longest growing seasons—were located and mapped. Winter cattle and sheep grazing areas were predicted to be associated with permanent log structures (ranch headquarters), and summer grazing sites (temporary camps) were predicted to be associated with limited scatters of historical artifacts and perhaps with less-permanent structures (e.g., simple corrals).

The model provides insight into the probable distribution of sites created by stockraising activities, and it provides a framework for assessing the significance of such sites. Although its spatial resolution is low, it does provide a way of estimating site presence for every 900 ha area, and it illustrates that the sites tend to be near creeks. It could be argued, with some justification, that the model is overly simplistic. This project should be recognized, however, as one of the earliest attempts to deal with Euroamerican ranch sites as a resource of concern to cultural resource managers and as a potential data base for acquiring important information about regional history. Viewed from that perspective, the model was successful. This model and the one developed by Hackenberger (1984; see below) have a similar procedural logic, and both were an outgrowth of a Forest Service reconnaissance/predictive modeling project (Knudson et al. 1982).

Bisti–Star Lake Project *Archaeological Variability within the Bisti–Star Lake Region.* Meade F. Kemrer, editor. ESCA-Tech. 1982

Archaeological investigations for the Bisti–Star Lake Project were funded by the Bureau of Land Management and carried out by personnel representing the Albuquerque office of ESCA-Tech, an environmental consulting firm. The modeling objectives for the project were to develop and refine methods capable of predicting the presence of sites with specific cultural and temporal characteristics. That information would then be used to generate formal predictions concerning the density of sites of various types throughout the project area (Kemrer 1982).

Sagebrush, rabbitbrush, greasewood, and other semiarid vegetation is characteristic of the dissected plateaus in the 31,413 ha project area, which lies within the San Juan Basin of New Mexico. Landsat data were generated and coded for the project area in 2 by 2 km grid units (400 ha each). Seventy-two environmental variables, consisting of different combinations of eight environmental classes (e.g., Avalon-Sheppard-Shiprock soil association and major washes), were derived from Landsat data; one data set (presence/absence of variable states) contained all the unique two-way interactions between environmental classes. The archaeological data base for the initial model consisted of existing site-file data from surveyed areas within and adjacent to the project area. Site type and content data as well as information on cultural/temporal affiliation were examined for more than 450 components. Eight site classes were developed using analysis of variance techniques. A backward step-wise multiple regression was used to reduce the number of environmental variables, and other linear equations were used for modeling site component densities. Projected site densities for the 2 by 2 km grid units were plotted on maps.

The project area was then subdivided into a number of leases, and a sample totaling about 4600 ha (ca. 15 percent of the total project area) was judgmentally selected and surveyed. Choice of parcels to be included in the judgmental sample was based, in part, on land ownership, size of sample units, and predicted cultural resource variability. A total of 92 sites and 213 isolated finds were documented. Some Paleoindian and Archaic sites were found (11 of 319 components), but most remains were classified as Anasazi, Navajo, historical, or lithic scatter sites. Resulting data were added to the existing site-file data base as a means of testing and refining the initial model. A regression analysis approach was again used to produce the refined model. When the augmented cultural resource data base was analyzed with 34 environmental variables, figures showing the percentage of explained variance were generated fot each of the site types. Mean site-frequency predictions were generated for more than 800 grid units and plotted on eight maps, one for each of the following site types: lithic sites, Anasazi sites, pre-1933 Navajo sites, post-1933 Navajo sites, total Navajo sites, and total sites.

The overall modeling approach yielded information on the range of variability in cultural/temporal components, site types, and site densities. The means by which this was accomplished and the overall reliability of the results are not always obvious. Much of the discussion on model development is difficult to comprehend, and decisions about selection of areas for survey were highly judgmental. The project area, the area from which the environmental data were extracted, and the survey area were all different, and the size of survey units differed from subarea to subarea. These factors may have affected the results of the statistical analysis.

There are also potential problems with the manner in which field information was gathered and analyzed. These problems make it difficult to replicate the overall approach and may well have caused the model to yield arbitrary results. Isolated finds, for example, were excluded from site density estimates. Unfortunately the criteria used to distinguish isolated finds from sites were not rigorous. In fact,

considerable overlap is likely given that different survey teams and different individuals operationalized the site and isolated find definitions:

> Sites were differentiated from isolated cultural occurrences on the basis of information potential. A site was defined as a locus manifesting the outcomes of past human behavior which contained more identifiable or potential scientific data values than could be effectively extracted at the time of survey. Isolated occurrences were defined as those cultural manifestations whose scientific data valnes could be adequately documented by the survey [Cella 1982:75].

Another factor that might have led to arbitrary results has to do with the manner in which sites were classified as to type. The most obvious case is the merging of identified Paleoindian and Archaic components with unidentified lithic components to create a single type. That procedure probably masks a significant portion of the observed cultural/temporal and site type variability, yet detection of that variability was one of the major goals of the project.

Ozark-St. Francis National Forests Project *A Cultural Resources Overview of the Ozark-St. Francis National Forests, Arkansas.* George Sabo III, B. Waddell, and J. H. House. Arkansas Archaeological Survey. 1982

Arkansas Archaeological Survey personnel conducted this overview project in the Ozark-St. Francis National Forests for the Forest Service (Sabo et al. 1982). The principal objectives were to assess the potential nature and distribution of prehistoric and historical sites in unsurveyed areas and to provide predictions concerning the nature and distribution of cultural resources. This information was to be incorporated into multiple resource management plans.

The 461,000 ha project area encompasses two national forests in northwestern and east-central Arkansas. As a means of generating expectations for the nature and distribution of cultural resources, a series of deductive adaptational models were developed. Four temporal periods were defined jointly by adaptation type and paleoenvironmental type: Late Pleistocene/Early Holocene hunting and gathering; Middle Holocene hunting and gathering; Late Holocene (post-Hypsithermal) hunting, gathering, and plant husbandry; and Late Holocene horticultural, hunting, and gathering. Initial narrative predictions were made concerning the distribution, content, and types of sites within each of four major environmental zones: river bottomland, upland slopes, bluff lines, and upland plateaus. A similar approach was used to define seven major and seven supplementary historical adaptation-type models. Examples of these ethnohistorically and historically recorded types include Osage (AD ?-1804), Creek (1794-1828), Spanish (1673-1803), pioneer hunter/herder (1803-ca. 1840), Civil War (1860-1875), resorts (ca. 1860-present), and Forest Service (1908-present).

Biophysical data, including elevation, soil types, topographic settings, physiographic subdivision, and vegetation types, were coded for 259 known sites that could be plotted reliably on USGS quadrangles. Q-mode cluster analyses were

performed separately on prehistoric and historical sites. Univariate and bivariate statistical procedures were used to determine which variables correlated best with site locations. The important variables were topographic setting, soil capability, distance to water, and elevation. The resulting inductive models yielded four clusters. These were qualitatively compared with expectations derived from the adaptation-type models. It was concluded that the inductive, Q-mode analyses generally supported the deductive models. Site likelihood zones based on topographic setting for historical and prehistoric sites were plotted on maps, and generalized site-density-potential values (high to low) were assigned to each zone.

As is the case with most deductive predictive modeling approaches, the end result of this project provides only limited spatial resolution for the predictions. In this case, most of the zones comprise thousands of hectares, and available data do not permit a finer resolution of density and/or potential anywhere within a given zone. Furthermore, this kind of model is difficult to falsify, largely because of its low spatial resolution and generalized treatment of site content data. It does, however, meet its objective in that predictions are made for the potential nature and distribution of cultural resources. The approach also allows for, and in fact encourages, the discovery of site types that are undocumented but theoretically expected in the study area. Examples include most of the Pleistocene site types and types representative of seventeenth-and eighteenth-century adaptations. Furthermore, the issue of site significance is divorced from the concept of site likelihood zones: the authors note that "significance must be determined on a case-by-case basis . . ., and a site in any likelihood zone could easily turn out to be highly significant" (Sabo et al. 1982:188).

Passaic River Basin Project *A Preliminary Cultural Resource Sensitivity Analysis for the Proposed Flood Control Facilities Construction in the Passaic River Basin of New Jersey.* Robert Hasenstab. Soil Systems, Inc. 1983

The New York District Corps of Engineers funded the Passaic River Project; Robert Hasenstab (University of Massachusetts, Amherst) implemented the project through a subcontract with Soils Systems, Inc., an environmental consulting firm based in Marietta, Georgia. The project's objectives were to estimate the quantities of cultural materials likely to be affected by proposed flood-control facilities and to define areas with a high probability of site occurrence (Hasenstab 1983).

The 1619 ha project area extends 160 linear km along the Passiac River, cross-cutting ridge and valley, piedmont, coastal plain, and tidal/estuarine areas, Urban and commercial developments occupy most of the impact zone, but 42 percent is either agricultural, forested, or classified as wetlands. The project area was subdivided into a high-resolution grid of 0.47 ha units (pixels) for which various environmental variables were coded; all manipulation and mapping utilized a GIS. Univariate statistical tests were employed to determine which environmental variables were most useful for their power to "retrodict" known site locations. Significant variables were found to be soil drainage, distance to nearest river, distance to minor

tributary confluence, and distance to a major tributary/river confluence. Grid cells were assigned a sensitivity rating by summarizing the various cultural component-variable ratings. The sensitivity models were then tested and revised using data derived from a survey of 300 pixels (ca. 140 ha) representing a stratified random sample of the project area (with some modifications). Overall, the sample fraction was about 6.5 percent of the impact zone. The survey techniques included limited but systematic subsurface testing within judgmentally selected pixels. Twenty-eight historical sites and 16 prehistoric sites were recorded. A series of computer-generated maps illustrated the final model on a pixel-by-pixel basis in terms of prehistoric archaeological sensitivity (high, medium, or low, based on the cultural component-variable ratings) and a combination of historical and prehistoric sensitivity.

The author concludes that the GIS approach "has greatly enhanced the capabilities for archaeological prediction and land-use management, . . . [but it] cannot be taken as a final solution to all cultural resource management problems" (Hasenstab 1983:13). The land managers did learn something new about the distribution of sites, but not much about their nature. Hasenstab's (1983:i–ii, 14–16) self-critique warrants close attention, since the problems he identifies are shared by many models: *(a)* the grid resolution may have been too coarse to detect important variables (such as small sandy knolls), *(b)* no attempt was made to deal with problems of spatial autocorrelation, *(c)* no consideration was given to understanding the effects of different variables on different site types, and *(d)* the fieldwork was probably not of sufficient scope to assess the model adequately. The approach is also problematic because it lumps together all prehistoric sites and thus tends to obscure the variability that is represented by thousands of years of human occupation.

Like some of the other models discussed here, this one also equates high likelihood zones with a high potential for the occurrence of significant sites. Furthermore, it equates low sensitivity with nonsignificance and with a lack of necessity for legal protection. This is demonstrated in the following statements from a subsection of the report entitled "Synthesis of Cultural Resources Sensitivity":

> Finally, 20 percent of the project area could be "written-off" legitimately. The low historic/low prehistoric sensitivity stratum (10 percent of the project area) would yield a very low return on encountered cultural resources. The medium historic/ low prehistoric sensitivity stratum (10 percent), as mentioned above, could be sacrificed, as a substantial portion of the medium sensitivity stratum will already have been sampled [Hasenstab 1983:134].

Such a conclusion does not seem compatible with a *preliminary* cultutal resource sensitivity analysis, which the title of the report indicates that this was intended to be. Neither does it seem to be compatible with the author's recognition that the sample survey may not have been of sufficient scope to permit adequate assessment of the model.

Grand Junction Resource Area Project A Manual for Predictive Site Location Models. Kenneth L. Kvamme. Draft report submitted to the Bureau of Land Management, Grand Junction District. 1983

The BLM funded the Grand Junction Resource Area Project as an overview of statistical classification procedures for predicting archaeological site locations. This summary emphasizes aspects of the project related to the development and testing of models in the Grand Junetion Resource Area. For that area, the objective was to develop quantitative models that could be used to predict likely locations of prehistoric sites (Kvamme 1983; see also Chapters 7, 8, and 10).

The project area encompasses some 438,996 ha of western Colorado uplands. Vegetation types characteristic of the area include desert grasslands as well as piñon-juniper woodlands and spruce-fir forests. The subareas of the district were stratified into five major biotic communities considered to occur in significant proportions across the landscape. A stratified proportional random sample of 65 ha quadrats (quarter sections) was selected from the physiographically defined subareas. One hundred quadrats were selected for survey, specifically to provide the data base for generating the models. The surveyed area amounted to about 1½ percent of the project area. Environmental data wete coded for site and nonsite locations. Through a series of statistical analyses, the following variables were found to be important in distinguishing between site and nonsite locations: biotic zone, vertical distance to permanent water, vantage point distance, slope, view, exposure, shelter within 100 m, and shelter within 250 m. The models were developed through a pattern-recognition approach using various multivariate analyses as classification tools, the most successful of which was logistic regression. Depending upon the particular approach used, GIS-based probability surface maps were generated to provide/illustrate predictions for sites and siteless loci in unsurveyed areas covering from 0.6 to 1 to 25 ha. The accuracy of the various models was tested independently using site-file and nonsite data, as well as split sampling techniques.

Kvamme's approach to predictive locational modeling is statistically and computationally more sophisticated than that exhibited by other projects summarized here. The report is clearly an important contribution in that it provides a thorough overview and many examples of a wide variety of statistical approaches to developing and testing inductive, or correlative, models. The project did not, however, achieve the goal stated by the author, namely "to model the locations of all sites, regardless of type, because all sites are of potential interest to Cultural Resource Management" (Kvamme 1983:69).

This suggestion that the Grand Junction Resource Area report failed to model the location of all sites is based on three observations. The first concerns the apparent paucity of sites in 38 percent of the project area. Kvamme suggests that the low density of sites (four were known) in the high-elevation community (which comprises 15 percent of the resource area) is a result, in part, of "the dense vegetational cover occurring at high elevations which inhibited site discovery" (1983:62). At the same time, only a few sites (26) occur in the desert community, which represents about 24 percent of the resource area (no explanation is offered for

this low density). Since 85 percent of the sites are in the piñon-juniper community, which constitutes only 62 percent of the project area, it was determined that

> Because of the paucity of sites in all but the pinyon-juniper communities, it will not be possible to make meaningful comparisons of site location patterning between communities in the analyses that follow although this was originally intended [Kvamme 1983:62].

Because two major zones with very different resource potentials for aboriginal hunters and gatherers were effectively excluded, it seems likely that potentially important site types were not modeled accurately.

Another aspect of this research that hampered modeling of all site locations was the exclusion of rockshelters from the analysis owing to the assumption that "their locations cannot be predicted because of the idiosyncratic geological processes that regulate their presence" (Kvamme 1983:68). Although identified rockshelters represent only 2.5 percent of the recorded sites, they have considerable potential to yield important information.

A final point concerns the arbitrary distinction drawn between sites (10 or more artifacts in a 20 m diameter area) and isolated occurrences (fewer than 10 artifacts in an area of the same size).

> In order to make the analysis of site locational patterning more manageable and also to reduce the idiosyncratic locational variation undoubtedly exhibited by isolated occurrences of artifacts (in many instances), only "concentrations" of artifacts were recorded as sites and analyzed here [Kvamme 1983:67].

Many archaeologists might argue that sites are often represented by fewer than 10 pieces of pottery or chipped, ground, or battered stone. Another potentially important site type—small, low artifact density—was therefore excluded from the model.

Kaibab and Cuba Study Area Projects *Theory and Model Building: Defining Survey Strategies for Locating Prehistoric Heritage Resources.* Linda S. Cordell and Dee F. Green, editors. Cultural Resources Document No. 3. Forest Service, Southwestern Regional Office. 1983

The Kaibab and Cuba study areas are part of a project sponsored by the Forest Service as a collaborative effort among archaeologists from academic and federal communities (Cordell and Green 1983). Specifically, participants in the endeavor were asked to formulate trial predictive models that could be refined and tested. The information summarized here is from two model-building articles, one about the Tusayan District in the Kaibab National Forest (Plog 1983a) and one about the Cuba District in the Santa Fe National Forest (Plog 1983b).

Study Area 1, the Tusayan Ranger District of the Kaibab National Forest, is located in northern Arizona. It is on an upland plateau that is dissected by

intermittent streams and covered by piñon-juniper, ponderosa, and sagebrush vegetation communities. Known sites in the area are the results of Archaic, Anasazi, and Cohonina occupations. The objectives for the Kaibab study area were to use previously derived information from a 1 percent sample survey (designed for planning purposes) to make predictions about site densities across the landscape and to "test" the predictions by comparing them with the results of intensive surveys conducted in nearby areas.

For the sample survey the 4858 ha study area was divided into zones based on drainage basins and vegetation types. Using the results of the 1 percent sample survey, the researcher estimated site densities for the various zones. The estimated densities were found to differ considerably from the observed densities in nearby intensively surveyed areas. The differences were judged to be the result of the nonquantitative fashion in which the estimated density figures were generated (e.g., there was no rationale for dividing the area into drainage basins, and zones without sample data were assigned zero density values). In the case of this trial formulation, it was concluded that "had SYMAP or some other spatial smoothing program been employed, a successful predictive model might have been generated" (Plog 1983a:66).

The Cuba District study area (Study Area 4) is a 3427 ha block unit in the forested upland zone of north-central New Mexico. In this case the objective was to examine the feasibility of doing predictive modeling by drawing upon the results of intensive surveys of the block area. The study area was surveyed in part by a Forest Service crew and in part by a contractor's crew. A total of 142 sites, all dating to the Gallina phase (AD 1150–1250), were documented. These included sites with surface structures, pithouses, towers, and check dams. An analysis of the survey data revealed that 96 percent of the sites were located on ridge tops, while this topographic feature constituted only 23 percent of the survey area. Even though few sites were found on the valley floors (and all of these were found by a single crew), it was recognized that these sites could potentially provide "important and unique evidence" about the area.

The researcher concluded, therefore, that if surveys in this study area were focused on the valley floors and ridge tops, coverage could be limited to 38 percent of the study area and almost all the sites would still be discovered. It was also argued that once a number of valley floors had been surveyed it should soon be possible to distinguish the characteristics of those valley floor ecosystems that would have associated sites from those that would not have sites. The author concludes his study by stating that "data from this study area result in as clear a definition of an approach for finding all sites with less than inventory survey as one can imagine" (Plog 1983b:78).

Both trial formulations of predictive models are presented in a brief and simple fashion. The models are mapped to illustrate the locations of high site-density zones within the outlines of the study areas. The lack of background information about these study area projects makes it difficult to understand how data were gathered. Much of the information necessary to compare this approach with others is not

readily obtainable from the report. It appears that the underlying purpose of this project was to determine whether or not portions of the areas could be exempted from on-the-ground survey by relying on the results of previous surveys in similar environmental settings. In the ease of the Kaibab area, for example, it is argued that had the appropriate "spatial smoothing program" been used, "rhe predictive model generated in the planning document would have allowed a no-survey decision to be made" (Plog 1983a:65).

What remains unexamined in these trial formulations is the reliability of the existing survey data. The problems attributed to a "nonquantitative" approach in the Kaibab study area can be alternatively explained by arguing that different people conducted the surveys for different reasons and that, consequently, the results are likely to be different. The uncritical acceptance of the survey results in the Cuba area—which indicated that throughout prehistory the area was inhabited only for a 100-year period, between AD 1150 and 1250—is also questionable. Could past erosional conditions have filled the valley floors so thar only relatively recent sediments are exposed, thus masking evidence that the area was also used or occupied by other groups of people? Is it possible that ground cover obscured all but the most obvious (i.e., architectural) culrural features? The information that one survey team found all the recorded valley floor sites indicates the potential for problems in data reliability; other things being equal, one might logically conclude that different survey methods were used. What may be needed here is not merely refinement and testing of trial formulations, but a reformulation of the approach to predictive modeling, one that recognizes the complex variation inherent in the archaeological record.

Fort Benning 2200-Acre Survey Project *An Intensive Survey of a 2,200 Acre Tract within a Proposed Maneuver Area at the Fort Benning Military Reservation.* P. M. Thomas, Jr., L. J. Campbell, M. T. Swanson, J. H. Altschul, and C. S. Weed. Report of Investigations No. 71. New World Research, Inc. 1983

The Department of Defense (U.S. Army Infantry Center and Fort Benning Military Reservation) funded this project, the second project carried out within the confines of Fort Benning to be summarized in this appendix. This study was administered by the Archeological Service Branch, Division of National Register Programs, National Park Service, Southeast Region (Atlanta) and carried out by personnel representing New World Research, an archaeological consulting firm based in Pollack, Louisiana. This project was designed to conduct an intensive survey and to test and refine a predictive model developed for the area three years earlier by another consulting firm (Kohler et al. 1980; see above). Information presented in this summary is from Thomas et al. (1983).

Pine forests, oak and oak/hickory uplands, bottomland hardwoods, wooded swamps, and mixed pine/hardwood forests are characteristic of the 8907 ha proposed maneuver area that was the focus of this project. A block of land amounting to about 10 percent of the project area (ca. 891 ha) was preselected and surveyed to provide a data base for evaluation of the larger maneuver area and for the testing of the

existing predictive model. Thirty-seven sites were identified: 20 prehistoric, 15 historical, and 2 with both prehistoric and historical components. Site locations were assessed according to a predictive model based on soil type, slope, and distance to water, which was developed by Kohler et al. (1980). The model was found to be basically sound but in need of some refinements, including a more accurate mapping of the distribution of soil types. In an effort to refine the model and determine which variables best explained the observed variation, a discriminant analysis was undertaken. Data for 10 environmental variables, including some information from a hypothetical catchment area with a 225 m radius, were coded at the 37 site locations and at 40 siteless locations. Ultimately, combinations of the variables were identified that could be used to define very high, high, low, and very low probabilities for encountering prehistoric and/or historical sites at any given location. A second discriminant analysis was performed on a data set from other portions of the project area; this data set consisted of 207 known sites and siteless points, including the 77 cases from the surveyed area. The discriminant analysis successfully reclassified more than 96 percent of the cases.

The project achieved its stated goals of testing and refining the existing predictive model. The refinements took the form of more accurate mapping of soil types and of the generation of a discriminant function that permits calculation of the probability of encountering a site at any given point on the landscape. Although a very low site-density zone is defined, it is neither tied to any significance determination nor used as an argument to exclude the area from future surveys. Like many of the other correlative or inductive models, this one masks much of the important variability in the archaeological record by lumping all prehistoric site types into one group.

Cisco Desert Project *A Class II Survey and Predictive Model of Selected Areas in the Cisco Desert, Grand County, Utah.* J. E. Bradley, W. R. Killian, G. R. Burns, and M. A. Martorano. Cultural Resources Report No. 10. Goodson & Associates. 1984

Goodson and Associates, a private consulting firm, conducted the Cisco Desert Project for the Bureau of Land Management. The project's modeling objectives were to use existing data to construct a predictive model for the location of site and siteless areas and to test the model with results of a sample survey (Bradley et al. 1984).

The 32,389 ha project area lies within the Colorado Plateau region of east-central Utah and is characterized by desert shrub, greasewood, and juniper woodland vegetation communities. Although the plan was to test an existing model, it soon became apparent that the existing model was inadequate for the project area, both because the project area had a much higher site density and because sites were found in many microenvironmental zones (e.g., dunes and rockshelters) that were not present in the areas for which the original model had been developed. The solution adopted was to build a model using information from a 5 percent sample survey conducted as part of the project, and then to test the model on data collected

during previous surveys. One hundred 16 ha (40 acre) tracts were selected for survey using a simple random sampling technique; an additional 17 tracts were selected on what amounted to a judgmental basis. A total of 126 sites were recorded within the randomly sampled 1619 ha area; 15 sites were historical and 111 were aboriginal, representing early Archaic through protohistoric occupation of the area. Eighty-eight sites (40 lithic scatters and 48 campsites) and 51 siteless locations from within the 5 percent sample were employed to construct four discriminant analysis models (two for each site type) using either "traditional" modeling variables, such as slope, distance to water, vegetation, etc., or soil unit variables. The soil unit models were found to be more accurate and easier to use. Sensitivity ratings for high, medium, low, and unknown (the sample for one soil type was too small for predictive purposes) chances of encountering a site were calculated on the basis of the various soil units. Soil unit/projected site density values were mapped for the entire project area. The overall results were judged to compare favorably with those generated from an existing model derived from a 10 percent sample survey of 58,705 ha in adjacent areas.

Although the manner in which the models were developed and tested differed from the original plan, the overall objective was achieved. More specifically, an environmental variable—soil unit—was identified as an accurate predictor of site locations, and areas of low site density were delineated for management purposes. An obvious shortcoming, however, is what the authors refer to as the lack of an adequate data base for making predictions in Soil Unit 9, which constitutes 7.8 percent of the project area. Too few transects were surveyed in areas with this soil unit, and too few sites were discovered in those transects to permit confident inferential model construction.

Bradley et al. (1984:88) draw the reader's attention to the fact that many of the sites misclassified in the discriminant analysis (ca. 15 percent) were in Soil Unit 2 (48.3 percent of the area and 0.95 sites per mi² in surveyed areas). Many of these sites were also located within 1287 m (0.8 mi) of an area of Soil Unit 3 (13.4 percent of the area, 27.35 sites per mi²). Given this situation, their recommendation with regard to additional survey of Soil Unit 2 areas is as follows:

> If survey requirements in this zone are waived by the BLM, isolated eligible sites may be endangered. It is recommended that all areas within .8 mile of soil units 3, 5, 8, and 9 continue to be surveyed in order to protect these sites and further test the model's accuracy. This .8 mile buffer includes sites misclassified by the soils model [Bradley et al. 1984:96].

Continued survey in the buffer zone would test the model only in regard to site density in the buffer zones; it would not be a test of whether National Register-eligible sites are present in the other portions of Soil Unit 2. This approach accepts the possible loss of an unknown number of sites in approximately 25–30 percent of the project area, and it recognizes that some of the sites may be eligible for inclusion in the National Register of Historic Places. By its reliance on modern environmental distributions, it potentially jeopardizes the opportunity to discover and investigate

sites that may have been utilized and/or occupied at times in the distant past when the desert shrub stratum, including Soil Unit 2, was more like today's juniper stratum (i.e., Soil Unit 8).

Route 13 Relief Corridor Project *A Cultural Resources Reconnaissance Planning Study of the Proposed Rt. 13 Relief Corridor, New Castle and Kent Counties, Delaware.* J. F. Custer, P. Jehle, T. Klatka, and T. Eveleigh. University of Delaware. 1984

The Route 13 project was funded by the Delaware Department of Transportation with the objective of identifying zones within a proposed highway corridor that were likely to contain significant prehistoric and/or historical resources. The project was conducted as an overview/planning study by personnel representing the Center for Archaeological Research at the University of Delaware (Custer et al. 1984).

Wetlands, agricultural lands, and urban areas occupy most of the 64.4 by 11.3 km project area (ca. 72,772 ha) in north-central Delaware. The predictive model was developed within different contexts: one for the environment and the other for regional cultural history. It relies heavily on the results of previous overviews. A number of site types (e.g., macroband basecamps, procurement sites, quarry sites, and industrial, commercial, and transportation sites) were recognized for various prehistoric and historical periods. Site types were characterized according to their environmental settings, and the information was summarized in tables that represent a general locational model. The general model was compared in a narrative with the results of a Landsat/Odessa terrain analysis (pixel size = 2.3 ha) that incorporated site locational information. Logistic regression analysis was used to correlate environmental zones with site presence. Maps were produced to illustrate known site locations and probability zones for different ages and kinds of prehistoric sites. Tables provide information about the relative potential for encountering significant historical sites in individual pixels. A separate and very general deductive model was developed to predict and explain the distribution of Adena mortuary/exchange sites. A second series of maps was generated to illustrate the high, moderate, and low sensitivity zones in terms of their potential for containing significant sites. In essence, high probability zones had the greatest sensitivity and the greatest potential for containing significant sites.

This project considers a wide range of site types in terms of their predicted locations and potential significance. The concept of significance is defined in a manner such that small, disturbed, and plow zone sites are largely excluded. Considerable attention is given to an assessment of the quality (i.e., reliability) of the information in existing site files. For most zones the available information is rated as "poor" or "fair." Given that kind of data the value of developing a series of correlative models for a wide variety of prehistoric and historical site types seems questionable. High probability zones and/or big sites with large quantities of artifacts are viewed as potential National Register properties and small procurement sites, as well as plow zone sites in general, are considered "not likely" to be eligible. The authors clearly state that their assessments are preliminary, however.

They also note that the data presented should not be viewed as a substitute for site location/identification surveys anywhere within the project area. Although "no specific fieldwork was carried out as part of this study" (Custer et al. 1984:1), some of the predictions made in the study were apparently field tested in 1984 and 1985. Results from this recent work were not included in the present review effort, but according to Custer (n.d.), "field tests of the predictions showed a 90 percent accuracy rate."

In general, the project fulfills its objective in that it succeeds in identifying zones that are likely to contain significant sites. The connecting links among regional prehistory/history, the existing data base, and the predictive models are difficult to follow, however, owing to the somewhat disorganized nature of the report.

Montane Hunter-Gatherer Project *Cultural Ecology and Economic Decision Making of Montane Hunter-Gatherers in Central Idaho.* Steven Hackenberger. M.A. thesis, Department of Anthropology, Washington State University. 1984

The Montane Hunter-Gather Project is a master's thesis submitted to Washington State University. It was developed with the objectives of *(a)* determining how well proportional resource use by montane hunter-gatherers could be predicted by comparing hypothetical decision-making strategies with observed resource distributions, and *(b)* determining whether archaeological data could be used to address the problem. The work is a by-product of a 1978 reconnaissance survey/model-building project (Knudson et al. 1982) funded in part by the Forest Service and the Idaho State Historical Society. Information presented here is from Hackenberger (1984).

The 1,216,800 ha project area is drained by the Middle Fork of the Salmon River and can be characterized as a forested montane environment with parklands and meadows. Environmental data—distribution of vegetation units, yields of browse vegetation, and distribution of plant, fish, and ungulate resources in terms of available calories for humans—were encoded for 520 23.3 ha (9 mi^2) grid units. These data were used to develop general predictions for hunter-gatherer settlement location, proportional resource use, and winter popularion aggregation. LaPlace, Savage, and Wald decision criteria were used in computer simulations to model long-term choices of site location based on resource density and yields. Ethnographic data provided analogs for modeling economic decision making among historical and late prehistoric occupants of the region.

These analyses indicated that models based on resource distributions or changes in distributions were more successful at predicting site location than models of various decision-making processes. Preliminary archaeological data were compared with predicted settlement locations and population sizes. Some of the predictions (e.g., locations of winter village sites) could be supported with available archaeological data, but in general, the researcher found that more survey would be required to provide data to test the models adequately.

This approach to predictive modeling, particularly the aspects that focus on monitoring distributions of food resources, is promising because it offers the potential for predicting and explaining the distribution of cultural resources. As the model now stands, however, its application is limited to the time periods for which ethnographic land-use data are available. Since spatial resolution is low and predictions are difficult to quantify the use of the approach is limited to the early planning stages of cultural resource management. The models are testable, however, and with refinement they could become more readily falsifiable. What is particularly promising about the approach is that predictive modeling for purposes of cultural resource management can be conducted in the context of problem-oriented investigations that are likely to yield information important in prehistory and history.

Tar Sands Project *The Tar Sands Project: Cultural Resource Inventory and Predictive Modeling in Central and Southern Utah.* Betsy L. Tipps. P-III Associates. 1984

The Tar Sands inventory/modeling project was funded by the Bureau of Land Management and carried out by individuals representing P-III Associates, an archaeological consulting firm based in Salt Lake City. The project's objectives included *(a)* implementation of a 5 percent inventory of each tract in the project area, *(b)* development of a site locational model that would correlate environmental characteristics with known site locations, *(c)* inventory of an additional 5 percent of each project area tract and use of the resulting data to test and refine the model, *(d)* development of projections of site density distributions and diversity of cultural resources based on the results of the 10 percent combined inventory, and *(e)* definition of the factors that determined cultural resource site selection and have explanatory value for predicting the location of sites. Information summarized here is from Tipps (1984).

The 69,635 ha study area lies in the Canyon Lands section of the Colorado Plateau and exhibits typical Great Basin vegetation patterns: shadscale, sagebrush, and piñon-juniper zones. Two 5 percent simple random samples (with some modification) of 65 ha quadrats were drawn for survey purposes from each of four large tracts. Including "buffer zones," some 7400 ha were surveyed and found to contain 155 sites (167 components) as well as a number of isolated finds. The sites represent occupations from the early Archaic to the historical periods. Prehistoric site density estimates with confidence intervals were made for each tract. Map-readable environmental variables were correlated with site locations in three of the four tracts using a discriminant analysis applied to data from one 5 percent sample. The results of the first analysis were tested and refined using the additional 5 percent sample data and a set of siteless areas. A final discriminant analysis was based on the 10 percent sample. Using six environmental variables—relief, elevation, distance to water, distance to nearest river, drainage, and quadrat vegetation cover—the analysis correctly classified 71 percent of the quadrats into categories of no sites, one site, and two or more sites; when these categories were combined, 93 percent of the quadrats with sites were classified correctly. Another predictive model was generated using Landsat imagery data and cluster analysis to classify the

area and provide probability estimates of site occurrence. Its utility for management purposes was found to be limited, however, because all the strata had similar probabilities of site occurrence.

This project achieved most of its goals, especially those related to the sample surveys and to finding correlations between environmental variables and site locations. In fact, this study represents one of the more sophisticated and better presented versions of the now-familiar correlative approach to predictive modeling (e.g., Larralde and Chandler 1981; Kemrer 1982; Kvamme 1983; Bradley et al. 1984). The discrimination of three classes of grid units—those with no sites, those with one site, and those with more than one site—may be an improvement over approaches that only distinguish between site-present and site-absent quadrats. It is also noteworthy that during the course of fieldwork an effort was made in some areas to determine whether there were buried cultural materials. Existing road cuts and cutbanks were examined, and a few buried sites were recorded. This practice seems advisable in areas noted for their long histories of high rates of erosion (e.g, the Southwest and the Great Basin).

One shortcoming of the discriminant and Landsat models was that the White Canyon tract was excluded from the analysis. This exclusion is unfortunate because even though this tract represents only 6.1 percent of the project area, it has an average density of 2.86 sites per quadrat. The discriminant model and the Landsat models are subject to other criticisms frequently made of projects using a correlative approach (see Berry 1984), including criticisms of arbitrary distinctions between sites and isolated finds.

The project was much less successful in achieving the goals of defining and explaining factors that determine site location. For example, the following partial explanation was offered for the success of the discriminant function in distinguishing quadrats with only one site:

> the single sites in these quadrats generally represent small, limited activity sites that occur in a localized anomalous portion of the quadrat. The quadrats in which these isolated sites are found may represent areas where more specialized or limited types of activities were occurring such as hunting or plant gathering or material procurement. For such sites variables such as distance to water, percent of quadrat cover, etc., may not be key factors in site location at all. We note, as do previous researchers, that site type is a critical factor in understanding the site selection process for prehistoric peoples [Tipps 1984:158].

These statements recognize the problem that lumping site types obscures important differences, but they do not explain why distance to water and vegetation type should be less useful for predicting the locations of hunting or vegetal procurement sites than for basecamps or other multiple activity sites. These explanations assume, as most correlation-based explanations do, that groups who used these anomalous quadrats (35.6 percent of all those with sites) for thousands of years all did so in essentially the same manner, in spite of significant changes in human population densities and technological developments, not to mention climatic

changes that surely affected the distribution of food resources. Even if such redundancy of land use could be demonstrated, it too would require considerable explanation.

Central Oregon Project Locating Significant Archaeological Sites by Landform Analysis in Centtal Oregon. Leslie E. Wildesen. Draft report submitted to the Bureau of Land Management, Oregon State Office and Prineville District Office. 1984

The central Oregon predictive modeling project was funded by the Bureau of Land Management and conducted by personnel representing Wildesen Associates, a Portland-based archaeological consulting firm. The project's objectives were to identify lands likely to contain significant prehistoric sites requiring "affirmative management action" and to identify lands not likely to retain an important archaeological record. The purpose of identifying these land categories was to focus efforts on sites that are subject to the requirements of the National Historic Preservation Act. Information summarized here was taken from a draft document by Wildesen (1984), which was circulated widely for review purposes.

The overall project encompasses an area comprising almost 1 million ha, of which 427,787 ha are managed by the BLM. Characteristic vegetation communities include sagebrush and grasslands as well as juniper and ponderosa forests. Within the larger area, 364 prehistoric sites were documented in existing site files, and 244 of these are on BLM land. All sites in the project area and all lands managed by the BLM were used to develop the model. The sites were judged to represent the full "functional and descriptive" range of site types known from the Desert West.

The concept of site significance was an important element of this study. Wildesen followed a previously established working definition for the concept of "important information," which was defined as

> substantive new information on northern Great Basin settlement or subsistence patterns, chronology, toolkits or technology, art, or intercultural relations (including travel or trade) [BLM 1982, cited in Wildesen 1984:2].

Wildesen (1984:3) goes on to note that, by implication, significant sites

> — will show evidence of more than one kind of use, or more than one use event;
>
> — will contain diagnostic tool types, comparable with existing typologies;
>
> — will exhibit physical integrity over more than 50 percent of their surface area;
>
> — may contain internally stratified sediments or cultural layers;
>
> — may contain artifacts or manufacturing debris, faunal remains, or constructed features (cairns, pits, painted or pecked rock art panels, or walls); or
>
> — may be related to similar or different sites within a specific geographic area (i.e., comprise part of a National Register District).

Ethnographic data were employed to identify the kinds of landforms used by Native Americans for various activities. Five landforms were identified as having been used ethnographically *and* as having the potential for containing sediments with "high physical integrity." These landforms are noted as having already yielded "archaeological sites with substantial scholarly values." These five landforms, along with two other landforms "known to contain archaeological remains of significant interest" (Wildesen 1984:4), were classified as *high probability landforms*. These seven landforms were calculated to represent only 7.3 percent of the project area. The model is presented in the form of text, tables, graphs, computer printouts, and maps that illustrate the locations of high probability landforms.

Explanations for and potential applications of the modeling approach were as follows:

> By focusing the analysis on where natural processes are not likely to have preserved intact archaeological evidence, as much as 93 percent of the study area can be removed from the potential data base. This does not mean that some evidence of prehistoric use may not be present on those acres, or that those acres were not used at some time in the past. It does mean that evidence of use is likely to be disturbed, inconclusive, or missing entirely from the record. Under such circumstances, it is very unlikely that the archaeological values of any sites located on these acres will warrant substantial archaeological resource management activity, or will require significant effort to resolve conflicts with other resource management activities [Wildesen 1984:5–6].

This project succeeded in identifying lands likely to contain significant prehistoric sites, but the methods used to accomplish these goals are problematic. First, there seems to have been no systematic attempt to evaluate the quality of data in the site files. If the Oregon site files are similar to those in orher parts of the United States, one mighr suspect that they need to be "cleaned" before being used ro construct models. Second, the model relies heavily on ethnographic analogy. Use of ethnographic information to define the areas habitually exploited by human groups for thousands of years and during different climatic regimes seems to be of limited value.

Of greater concern is the approach to defining "significant" sites. Categorical criteria are established for defining significance, and they virtually exclude small, disturbed, and plow zone sites, which have long been argued to be potentially significant (Talmage et al. 1977). Furthermore, the criteria do not acknowledge the potential for some site types (e.g., task-specific sires or small residential sites that might have been disturbed by natural processes) to contribute important information regarding significant research topics (e.g., land use systems for mid-Holocene hunter-gatherers). Removal of a 396,559 ha area encompassing an undetermined number of cultural resources from the potential data base may be premature, especially if this is done on the basis of existing, but unevaluated, survey data. The significance criteria outlined in Wildesen (1984) imply that significance is related directly to the degree to which an archaeological site can be considered to encapsulate an undisrorted view of the past. Binford has responded to those who share this

expectation by noting that "seeking a reconstructed Pompeii is an unrealistic and unprofitable goal in the light of knowledge we have and the data available to us in [the archaeological] record" (Binford 1981:206). It may be possible to construct a predictive model that can be used to "write off" areas because they contain only insignificant sites, but in the draft document summarized here, Wildesen (1984) does not present a convincing argument that the data base in question is adequate for this purpose.

CONCLUDING COMMENTS

This survey of predictive locational models is intended to present information on a range of approaches to predictive models in different areas and for different kinds of cultural resources. This appendix differs from the other sections of this volume in that it is a sample inventory of what has been and is being done in predictive modeling; it is not an evaluation of how predictive modeling is expected to be done or how it should be done in the future. The concluding paragraphs in the synopses of the projects are narrative assessments of how well the projects achieved stated objectives and, as such, are more judgmental than descriptive.

The goals of this survey were *(a)* to summarize projects representative of the known range of variation in approaches, geographic settings, and types of resources being modeled; *(b)* to provide a descriptive summary and assessment of the individual models; *(c)* to present data that facilitate comparisons among the different models; and *(d)* to provide enough information to permit the reader to make an independent assessment of the predictive locational modeling approaches reviewed. Although this survey was not designed to be a synthetic statement concerning predictive modeling, nor a critical review of individual projects, it does seem appropriate to end with a few comments of a more synthetic nature. Those offered here are based mainly on the detailed examination of these 22 project reports and on a perusal of many others.

The following discussion is intended to address two general questions. Do existing models contribute substantially to the management of potentially significant, nonrenewable cultural resources? And do they contribute information important to our understanding of history or prehistory? It is clear that some of the predictive models contribute information important to history or prehistory. Those with the potential for explaining aspects of human behavior are likely to be of special interest to archaeologists. Other predictive models provide probability estimates for encountering a particular kind of site at a specific place on the landscape, and that information is of special interest to land managers charged with protecting significant sites. None of the models assessed here have both explained significant aspects of human behavior *and* predicted the probability of finding evidence of specific behavioral patterns at specific places on the landscape, however.

Given the widespread perception that cultural resource management and research goals are separate and not especially compatible, this lack of models that meet both goals may not be surprising. It is not inevitable, however, because predictive modeling has the *potential* to contribute information important to both managing and understanding cultural resources. Granted that predictive modeling has not been perfected, what has it contributed during the past decade?

In the first place, more sites are being discovered and documented in a wider range of environmental settings than would have been the case 10 years ago. This is partially because sample surveys that provide the data base for predicting the total number of sites are often designed specifically to detect the range of site types in different settings. At the same time, there is an increased awareness that a high proportion of the extant archaeological materials is likely to be found in a small proportion of the landscape. Conversely, there is recognition of the potential that important cultural resources will be discovered within those portions of the landscape with lower site densities. Furthermore, it is becoming clear that there are few, if any, areas without any evidence of utilization by human groups. These contributions mean that cultural resource specialists, whether managers or archaeologists, are in a position to better understand the nature of cultural resources in a given area and the distribution of different kinds of archaeological materials on the landscape.

Development and use of predictive models also has focused attention on the interrelationships between environmental factors and site locations. The search for significant spatial correlations has identified many key environmental variables useful in predicting site locations. By knowing which environmental settings are likely to have certain kinds of sites, managers can determine how those areas can be managed with minimal effect on cultural resources. The correlations also provide data bases useful in assessing site function and testing models about land-use systems. Inclusion of information about *past* environmental settings is likely to be particularly useful in understanding how and why prehistoric groups used the landscape in a particular fashion.

Another contribution of predictive modeling has been the compilation of quantitative, as opposed to qualitative, data bases. With information on the estimated density and distribution of cultural resources, land managers can develop more effective plans for the long-term conservation of significant cultural resources. Given reliable survey methods and quantitative results, inter- and intraregional comparisons of site distributions can be made, along with comparisons of densities or other measures of the intensity of use. In turn, the data from these comparisons are useful in testing models about many aspects of past human behavior.

The predictive modeling approach has also resulted in a number of trends that may not contribute substantially to the acquisition of important information about history or prehistory. Some of the trends may actually hamper the well-informed management of nonrenewable cultural resources. Of potential concern are the models that provide probability estimates for encountering a generic site—one that could be of any type or age—at a particular point on the landscape. The generic site approach can imply that all sites are of equal importance, when clearly they are not.

Land managers must protect *only* the significant ones. This suggests the need to become more discriminating about what is being predicted.

Although predictive modeling has focused attention on the interrelationship between environmental factors and site locations, there is considerable variation among environmental variables that ostensibly predict site locations. Among the more common predictors are specific values for vegetation type, proximity to water, landform, solar exposure, soil type, slope, and elevation. Site locations and behavioral patterns that led to the deposition of materials probably correlate spatially with many other key environmental factors. Regrettably, the reader is often left with no information as to the significance and explanatory value of these correlations. The importance of correlations is manifested in their ability to predict site locations, especially those judged to be significant in terms of National Register criteria. In turn, site significance is determined by the resource's potential to contribute important information. That determination often requires understanding of why environmental variables correlate highly with site locations and/or with the kinds of human behavior that account for the site locations.

Identifying key environmental variables without explaining how and why they correlate with site location is tantamount to making predictions in a cultural and behavioral void. A review of the project summaries presented here illustrates a tendency to predict where sites should be found without adequately addressing the question of how humans used the environment. There is little discussion about relationships between the nature and distribution of basic food and nonfood resources on the one hand and complex human land-use systems on the other. A detailed study of some predictive models might convince the reader that the primary goal is to predict the distribution and density of prehistoric *things* on the landscape. Such predictions may be useful, but usually only in conjunction with other data that allow greater discrimination among the things predicted.

The tendency in many predictive models to avoid explanation and to make predictions in a cultural and behavioral void probably is related to a trend toward development and utilization of new technologies. Computers are the focal point of the new technologies because many of the modeling approaches depend on complex statistics and massive data files. GIS and Landsat are examples of new technologies that facilitate reliable point predictions. There is a danger, however, that these technologies could become the end product, rather than serving as a source of information useful in managing and understanding significant cultural resources. Given an emphasis on new technologies and the finite amount of time and money allocated to cultural resource management projects, there seems to be little time to study why the archaeological record appears as it does. Natural and cultural transformation processes are seldom discussed, and examination of human land-use systems is the exception rather than the rule in the predictive models reviewed here. It should be recognized, however, that the use of GIS, Landsat, and multivariate statistics is relatively new in predictive locational modeling. As with many new technologies, they can be expected to be used as a means to more informative ends as the science of predictive modeling matures.

In some cases there is an overreliance on current vegetation and young landforms to predict the occurrence of sites. An example would be the presence of sand dunes formed 4000 years ago as predictors of locations occupied by people 5000 years ago. Although these two events could be related, the underlying mechanisms are seldom discussed. Equally bewildering would be the significance of high positive correlations between the location of a 4000-year-old piñon-juniper forest and that of a hunting site occupied 6000 years ago, when the area may have been dominated by greasewood and sagebrush. It would seem more appropriate to identify environmental variables that are useful in predicting site locations *and* explaining the relationships.

Sand dunes, forests, and other aspects of the environment often act to bury or obscure cultural materials. Although this statement is an axiom to cultural resource specialists, most predictive models are not concerned with the discovery of buried or otherwise obscured sites. Discussions about depositional processes and the ages of landforms are seldom included in predictive models. In general, there is a paucity of discussions about the visibility of cultural materials on the surface, and discussions of survey methods rarely include a section on techniques used to find buried sites. Only a few of the models reviewed here address the relationship between the theoretically expected tange of site types and the range of site types recorded in the region or in specific survey areas. Fluvial and aeolian processes clearly act to bury older sites in many areas, and forest litter obscures hundreds of sites in other areas. If predictive modeling is designed to provide useful information on the distribution and density of all site types, the models should incorporate information on depositional and erosional processes and their effect on the archaeological record.

Another factor that limits the potential contributions of predictive modeling is an overreliance on the ethnographic record in predicting prehistoric site distributions. Investigators often assume that the settlement and subsistence patterns documented in the ethnographic record are manifested throughout the archaeological record. In other words, the investigators assume that by knowing something about settlement and subsistence patterns during the "ethnographic present" they also know where people camped and what they ate during the previous millennia. Detailed discussions of the time depth for the ethnographic pattern are uncommon. There are equally few in-depth discussions about the kinds of land-use systems that may have operated before human populations reached historical levels, or before they were decimated by European diseases, or before the density and distribution of large land mammals were reduced by environmental factors and/or human agents. Seen from this perspective, the ethnographic record may not provide information useful in predicting the locations of sites representative of land-use systems with very different settlement and subsistence patterns. In fact, overreliance on the ethnographic record is likely to inhibit detection of the range of site types present in the archaeological record.

Finally, there *may* be a growing tendency to "write off" large tracts of land by not recommending an inventory-level survey. Although the sample of models summarized in this appendix is not statistically representative of the universe of

predictive locational models, it is informative to note that about 23 percent of them include statements that either open the door to "writing off" large tracts of land or actually recommend it. None of the reports written prior to 1980 make such recommendations, but at least one report written that year makes that implication. Two such recommendations were made in 1983, and two others in 1984. Whether or not there is hard evidence for a growing tendency toward such recommendations is debatable, and in any case there may be justifications for some of those recommendations.

The decision to not recommend an inventory survey is usually made on the basis of sample survey data and/or information drawn from a review of available site-file data. Areas are usually written off because no sites are expected to occur there or because those that do occur there are not expected to be significant. The main problem with this procedure is that the reliability of the data base used for making the recommendation is usually questionable. The reliability of the data base depends upon the soundness of survey methods and/or upon the approach used to determine site significance. A second problem is that recommendations to write off an area without conducting an inventory survey tend to be based on the distribution of sites of known types, sites that were discovered using methods designed to find the best-known kinds of sites. This approach does not encourage the discovery of unknown but theoretically expected site types; rather, it focuses on refining established models. Generally, this encourages additional discoveries of sites of the best represented kinds at the expense of older sites and site types that are not readily visible on the surface. Exempting large areas of the landscape from inventory survey without assessing the reliability of the data base has the potential of ensuring that the range of site types remains undocumented.

The use of data generated by predictive locational models to legitimize no-survey recommendations is of particular concern because of the nature of cultural resources. Cultural resources are potentially important to many people for many different reasons, and they are *nonrenewable.* Once nonrenewable cultural resources are written off, they are likely to be excluded from further study, regardless of the validity of the rationale for this recommendation. In fact, the legality and ethics of writing off resources, especially on the basis of dubious data, is now being questioned. This is evidenced by lawsuits being brought against agencies that have cleared areas containing archaeological materials and by the increasing national dialog among archaeologists about this subject (Darsie and Keyser 1985; Tainter 1984).

Overall there is considerable variability in approaches among predictive models, both among those conducted in the systemic context and among those carried out in the analytic context. All of the models reviewed here were developed to provide information useful in the management of significant, nonrenewable cultural resources and/or information important to our understanding of history or prehistory. Although there are examples of models that provided information of special use to land managers and of models useful in explaining aspects of human behavior, none of the models assessed here were successful in providing both kinds of

information. Even so, it seems clear that predictive modeling, as used in cultural resource management, has the *potential* to provide both kinds of information. Given the relative recency of predictive locational modeling as a scientific approach in cultural resource management, *both uses and abuses* of it should be expected (Ambler 1984). Likewise, it should be anticipated that the potential to contribute a wide range of useful information will be realized as the science of predictive modeling matures. This volume was designed to provide the reader with information about how that potential might be realized.

Several people worked with me to bring this appendix to its present form, and I would like to acknowledge their assistance. Beth Miksa helped to compile the descriptive information summarized in the tables. Eileen Draper drafted the figure. Lorna Elliott willingly typed several versions of the draft manuscript and patiently formatted and reformatted the tables. Deborah Olson volunteered her assistance in proofing the manuscript. Tim Kohler constructively criticized earlier versions of the manuscript and helped me to clarify some of the points I was trying to make. Dan Martin and the other BLM and Forest Service personnel involved with the project, as well as the chapter authors and the editors, freely shared their ideas about predictive locational modeling and offered useful suggestions for how one might go about summarizing, comparing, and assessing the results of different modeling efforts. They also provided references for dozens of predictive locational models. I would especially like to thank the anonymous reviewers for pointing out inconsistencies in the review draft. Finally, June-el Piper and Lynne Sebastian merit special acknowledgment for their diligence and hard work in transforming my final draft into something significantly more presentable. Although those who assisted me deserve credit for their ideas and other contributions, I bear the responsibility for the contents of this appendix, including any errors it may contain.

REFERENCES CITED

Ambler, J. Richard

1984 The Use and Abuse of Predictive Modeling in Cultural Resources Management. *American Antiquity* 4:140–145.

Barber, Russell, and Michael E. Roberts

1979 *Archaeology and Paleontology.* Summary and Analysis of Cultural Resource Information on the Continental Shelf from the Bay of Fundy to Cape Hatteras, Final Report, vol. II. Institute for Conservation Archaeology, Peabody Museum, Harvard University. Contract No. AA-551-CT8-18. Submitted to the Bureau of Land Management. Copies available from the Institute for Conservation Archaeology, Cambridge, Massachusetts.

Berry, Michael S.

1984 Sampling and Predictive Modeling on Federal Lands in the West. *American Antiquity* 49:842–853.

Binford, Lewis R.

1981 Behavioral Archaeology and the "Pompeii Premise." *Journal of Anthropological Research* 37:195–208.

Bradley, J. E., W. R. Killian, G. R. Burns, and M. A. Martorano
1984 *A Class II Survey and Predictive Model of Selected Areas in the Cisco Desert, Grand County, Utah.* Cultural Resources Report No. 10. Goodson & Associates. Contract No. YA-553-CT2-138. Submitted to the Bureau of Land Management. Copies available from Goodson and Associates, Denver.

Bureau of Land Management (BLM)
1982 Brothers Grazing Management Program. Draft Environmental Impact Statement. Bureau of Land Management, Prineville District Office, Prineville, Oregon.

Cella, Nancy Hewett
1982 Field Methods. In *Archaeological Variability within the Bitti-Star Lake Region, Northwestern New Mexico,* edited by Meade F. Kemrer, pp. 74–75. ESCA-Tech. Contract No. YA-553-CTO-114 [SBA 580458(a) 80-C-0124]. Submitted to the Bureau of Land Management. Copies available from the Bureau of Land Management, Albuquerque District Office.

Cordell, Linda S., and Dee F. Green, editors
1983 *Theory and Model Building: Defining Survey Strategies for Locating Prehistoric Heritage Resources.* Cultural Resources Document No. 3. Forest Service, Southwestern Regional Office, Albuquerque.

Custer, Jay F.
n.d. Public Information Handout: U.S. Route 13 Corridor Study 1984–1985. Archaeology from 900 Kilometers up in the Sky. Copies available from the Delaware Department of Transportation, Division of Highways, Dover.

Custer, J. F., P. Jehle, T. Klatka, and T. Eveleigh
1984 *A Cultural Resources Reconnaissance Planning Study of the Proposed Rt. 13 Relief Corridor, New Castle and Kent Counties, Delaware.* Archaeology Series No. 30. University of Delaware. Delaware DOT Contract No. 83-110-01, FHWA Federal Aid Project F-1001(16). Submitted to the Delaware Department of Transportation and the U.S. Department of Transportation. Copies available from the University of Delaware, Newark.

Darsie, Richard F., and James D. Keyser, editors
1985 *Archaeological Inventory and Predictive Modelling in the Pacific Northwest.* Studies in Cultural Resources Management No. 6. Forest Service, Pacific Northwest Region.

Davis, C. W., D. C. Linck, K. M. Schoenberg, and H. M. Shields
1981 *Research Overview and Summary.* Slogging, Humping and Mucking Through NPR-A: An Archaeological Interlude, vol. 1. Anthropological and Historic Preservation Cooperative Park Studies Unit Occasional Paper No. 25. University of Alaska, Fairbanks.

DeBloois, Evan I.
1975 *The Elk Ridge Archaeological Project: A Test of Random Sampling in Archaeological Surveying.* Cultural Resources Report No. 2. USDA Forest Service, Intermountain Region.

Driver, Harold E.
1961 *Indians of North America.* University of Chicago Press, Chicago.

Ebert, James I., and Galen N. Brown
1981 Remote Sensing in the NPR-A Cultural Resources Assessment. In *Slogging, Humping and Mucking Through NPR-A: An Archaeological Interlude,* edited by C. W. Davis, D. C. Linck, K. M. Schoenberg, and H. M. Shields, pp. 349–419. Anthropological and Historic Preservation Cooperative Park Studies Unit Occasional Paper No. 25. University of Alaska, Fairbanks.

Hackenberger, Steven
1984 *Cultural Ecology and Economic Decision Making of Montane Hunter-Gatherers in Central Idaho.* M.A. thesis, Department of Anthropology, Washington State University, Pullman.

Hasenstab, Robert

1983 *A Preliminary Cultural Resource Sensitivity Analysis for the Proposed Flood Control Facilities Construction in the Passaic River Basin of New Jersey.* Submitted to Soil Systems, Inc., and the Passaic River Basin Special Studies Branch, Department of the Army. Copies available from Soil Systems, Inc., Marietta, Georgia.

Hunt, Charles B.

1974 *Natural Regions of the United States and Canada.* W. H. Freeman, San Francisco.

James, S. E., R. Knudson, A. Kane, and D. Breternitz

1983 Predicting Site Significance: A Management Application of High-Resolution Modeling. Paper presented at the 48th Annual Meeting of the Society for American Archaeology, Pittsburgh.

Jermann, Jerry V., and Stephen Aaberg

1976 *Archaeological Reconnaissance in the Libby Reservoir–Lake Koocanusa Area, Northwestern Montana.* Department of Sociology, Montana State University. Contract No. DACW67-76-SU-1519. Submitted to the Army Corps of Engineers, Seattle District. Copies available from the Montana State University, Bozeman.

Kemrer, Meade F., editor

1982 *Archaeological Variability within the Bisti–Star Lake Region, Northwestern New Mexico.* ESCA-Tech. Contract No. YA-553-CTO-114 [SBA 580458(a) 80-C-0124]. Submitted to the Bureau of Land Management. Copies available from the Bureau of Land Management, Albuquerque District Office.

Knudson, R., D. Stapp, S. Hackenberger, W. D. Lipe, and M. Rossillon

1982 *A Cultural Resource Reconnaissance in the Middle Fork Salmon River Basin, Idaho, 1978.* University of Idaho Anthropological Research Manuscript Series 67. Laboratory of Anthropology, University of Idaho, Moscow.

Kohler, Timothy A., and Sandra C. Parker

1986 Predictive Locational Modeling. In *Advances in Archaeological Method and Theory*, vol. 9, edited by Michael B. Schiffer, pp. 397–452. Academic Press, New York.

Kohler, T. A., T. P. Desjeans, C. Feiss, and D. E. Thompson

1980 *An Archaeological Survey of Selected Areas of the Fort Benning Military Reservation, Alabama and Georgia.* Remote Sensing Analysts. Contract No. C-5716(78). Submitted to the Heritage, Conservation and Recreation Service, Interagency Archeological Services, Atlanta. Copies available from the Interagency Archeological Services, Atlanta.

Kvamme, Kenneth L.

1983 A Manual for Predictive Site Location Models: Examples from the Grand Junction District, Colorado. Draft report submitted to the Bureau of Land Management, Grand Junction District, Colorado.

Larralde, Signa L., and Susan M. Chandler

1981 *Archaeological Inventory in the Seep Ridge Cultural Study Tract, Uintah County, Northeastern Utah, with a Regional Predictive Model for Site Location.* Nickens and Associates. Contract No. YA-553-CTO-1008. Submitted to the Bureau of Land Management. (Published in 1982 as Cultural Resource Series No. 11, Bureau of Land Management, Utah State Office, Salt Lake City)

Mierendorf, R. R., T. K. Eller, D. Carlevato, and P. A. McLeod

1981 *A Cultural Resources Predictive Land Use Model for the Okanogan Highlands: An Impact Assessment for the Bonneville Power Administration's Chief Joseph Dam–Tonaiket and Grand Coulee–Keller Transmission Lines.* Bonneville Cultural Resources Group Report No. 100-2. Contract Nos. DE-AC79-81BP21946 and DE-AC79-79BP14939. Submitted to Bonneville Power Administration, Portland, Oregon. Copies available from Eastern Washington University, Cheney.

Muto, Guy R., and Joel Gunn

1980 *A Study of Late Quaternary Environments and Early Man Along the Tombigbee River, Alabama and Mississippi, Phase I,* vols. I and II. Benham Blair and Affiliates. Contract No. CO-9002(79). Draft report submitted to the Heritage, Conservation and Recreation Service, Washington, D.C.; the Interagency Archeological Services, Atlanta; and the Army Corps of Engineers, South Atlantic District, Mobile. Copies available from Army Corps of Engineers, South Atlantic District, Mobile.

Plog, Fred

1983a Study Area 1: Kaibab National Forest. In *Theory and Model Building: Defining Survey Strategies for Locating Prehistoric Heritage Resources,* edited by Linda S. Cordell and Dee F. Green, pp. 63–66. Cultural Resources Document No. 3. Forest Service, Southwestern Regional Office, Albuquerque.

1983b Study Area 4: Cuba District, Santa Fe National Forest. In *Theory and Model Building: Defining Survey Strategies for Locating Prehistoric Heritage Resources,* edited by Linda S. Cordell and Dee F. Green, pp. 76–78. Cultural Resources Document No. 3. Forest Service, Southwestern Regional Office, Albuquerque.

Rossillon, Mary P.

1981 *An Overview History in the Drainage Basin of the Middle Fork of the Salmon River.* Cultural Resources Report No. 6. USDA Forest Service, Intermountain Region.

Sabo, George, III, D. B. Waddell, and J. H. House

1982 *A Cultural Resources Overview of the Ozark–St. Francis National Forests, Arkansas.* Arkansas Archaeological Survey. Project No. 497. Submitted to USDA Forest Service, Ozark–St. Francis National Forests. Copies available from Arkansas Archaeological Survey, Fayetteville.

Steward, Julian

1938 *Basin-Plateau Aboriginal Sociopolitical Groups.* Bureau of American Ethnology Bulletin 120. Washington, D.C.

Tainter, Joseph A., assembler and editor

1984 Predictive Modeling and the McKinley Mine Dilemma. *American Archaeology* 4:82–113.

Talmage, V., O. Chesler, and the staff of the Interagency Archeological Services

1977 *The Importance of Small, Surface, and Disturbed Sites as Sources of Significant Archaeological Data.* Cultural Resource Management Studies. National Park Service, Washington, D.C.

Thomas, David Hurst

1973 An Empirical Test for Steward's Model of Great Basin Settlement Patterns. *American Antiquity* 38:155–176.

1975 Nonsite Sampling in Archaeology: Up the Creek without a Site? In *Sampling in Archaeology,* edited by James W. Mueller, pp. 61–81. University of Arizona Press, Tucson.

Thomas, P. M., Jr., L. J. Campbell, M. T. Swanson, J. H. Altschul, and C. S. Weed

1983 *An Intensive Survey of a 2,200 Acre Tract within a Proposed Maneuver Area at the Fort Benning Military Reservation, Chattahoochee County, Georgia.* Report of Investigations No. 71. New World Research, Inc. Contract No. CX5000-2-0087. Submitted to Archeological Services Branch, Division of National Register Programs, National Park Service, Southeast Regional Office, Atlanta. Copies available from National Park Service, Atlanta.

Tipps, Betsy L.

1984 *The Tar Sands Project: Cultural Resource Inventory and Predictive Modeling in Central and Southern Utah.* P-III Associates. Contract No. YA551-CT3-340038. Submitted to Bureau of Land Management. Copies available from P-III Associates, Salt Lake City.

Wildesen, Leslie E.

1984 Locating Significant Archaeological Sites by Landform Analysis in Central Oregon. Draft report submitted to Bureau of Land Management, Oregon State Office and Prineville District Office.

Woodward-Clyde Consultants

1978 Appendix E. In *Well Field Development Plan for the Wasson Field–Denver Unit CO_2 Project Environmental Impact Report*. Draft report submitted to Shell Oil Corporation. Copies available from Woodward-Clyde Consultants, San Francisco.

INDEX

A

Absolute probability, 179

Accuracy, 16, 25, 38, 51, 52, 61, 65, 72, 75, 87, 142, 151, 205, 237, 238, 242, 243, 279, 282, 287, 288, 303, 308, 316, 327, 347, 355, 365, 371-374, 377, 378, 381, 389, 390, 392-396, 400, 402, 403, 414, 444, 506, 507, 516, 520, 531, 532, 556, 561, 562, 564-566, 569, 575, 576, 580, 583, 609, 615, 617, 624, 629, 631

Activity sets, 156, 157

Activity space, 4, 331, 382, 383
  definition of, 331

Adaptation type, 6, 40, 41, 50, 51, 67, 68, 138, 621

Advisory Council on Historic Preservation (ACHP), 11, 573

Agriculture, 8, 20, 31, 33, 81, 87, 112, 114, 132, 134, 135, 137, 141, 192, 258, 306, 345, 440, 466, 475, 478, 517, 518, 610, 614, 622, 630

American Southwest, 6, 31, 135, 276, 280

Ammerman, 19, 82, 105, 156, 157, 285

Analytic context, 37, 38, 42, 43, 52

Analytic units
  definition of, 286

Anasazi, 195, 581, 611, 620, 626

Antiquities Act, 34

Archaeological context, 3, 8, 37, 47, 82, 350, 382, 467, 577

Archaeological explanation
  definition of, 128

Archaeological record (see also Assemblage; Nonsite; Site)
  formation processes (see also Depositional processes; Postdepositional processes), 32, 51, 99, 100, 105, 109, 118, 143, 144, 146, 147, 155, 156, 158, 576, 579, 638

Archaic, 6, 82, 134, 258, 276, 470, 475, 481, 610, 611, 614, 620, 621, 626, 629, 632

Areal models, 63, 68, 69, 72, 74, 75, 175, 176, 258, 270, 272, 289, 293, 352, 470, 481, 552

Arizona, 20

Arkansas, 67, 220, 330, 387, 407, 518, 582, 621

Artifact collection (see No collection survey/policy)

Aspect variable, 358, 440, 479, 497, 516

Assemblage (see also Subsurface observations; Surface observations)
  coarse-grained, 105, 117, 121
  fine-grained, 121
  types/categories, 39, 43, 51, 116, 122, 124, 126, 146, 151, 154, 155, 157, 158, 206, 283, 288, 289, 475, 608
  variability, 116, 117, 119, 120, 121, 155, 206, 289

Associational models, 63, 66-68, 75, 85, 99

Assumptions,
  statistical, 16, 68, 74, 83, 85, 156, 173, 181-184, 186, 195-197, 199, 203, 204, 210-216, 221-224, 235, 251, 265, 269, 272, 273, 314, 327, 340, 351, 352, 355, 356, 358, 364, 365, 367, 371-373, 393, 394, 400, 402, 409, 415, 521, 560, 566, 568
  theoretical, 21, 28, 29, 31, 33, 34, 41, 46, 50, 51, 67, 72, 107, 108, 119, 126, 128, 146, 175, 200, 301, 329, 455, 456, 479, 481, 612, 625

Athabaskan, 6

Autocorrelation, 67, 69, 214, 352, 353, 357, 358, 394, 511, 521, 528, 565, 567, 623

B
Barber, 35, 46-48, 581, 611
Basin of Mexico, 20, 258, 259, 277, 281, 283
Basketmaker, 470, 609, 611
Base rate probability, 392
Bates, 29
Baumhoff, 30
Bayesian logic, 206, 235, 237, 411, 538
Bennett, 29
Bernoulli trial/distribution, 190, 191, 196
Bettinger, 29
BIA (see Bureau of Indian Affairs)
Binary response variable, 73, 207-209
Binford, 35, 44, 63, 79, 82, 102, 107-110, 112-114, 117-121, 123, 124, 130, 132, 133, 149, 157, 283, 294, 303, 467, 635, 636
Binomial distribution, 191-193, 216, 268, 270, 403
Biotic variables, 338, 624
Bivariate statistics, 174, 181, 200, 202, 251, 622
Black Mesa, 287
Blalock, 38, 39, 52, 193
BLM (see Bureau of Land Management)
Boas, 24, 25
Bodin, 22-24
Boyce, 20, 21
Brookfield, 29
Budgetary constraints (see also Cost), 304
Bureau of Indian Affairs (BIA), 450, 535
Bureau of Land Management (BLM), 9-11, 13, 34, 73, 75, 79, 86, 97, 122, 148, 152, 265, 270, 278, 279, 287, 306, 356, 394, 399, 402, 447, 535, 571, 573, 575, 579, 610, 611, 616, 619, 624, 628, 629, 632, 634
  District Offices, 279
Bureau of Reclamation, 148, 149, 274, 475

C
Cahokia, 258
Canonical
  correlation, 232
  discriminant functions, 222, 224, 227, 229, 231, 233, 235
Carr, 41, 42, 47, 51, 52, 123, 124, 155, 156, 158, 326, 338
Carrying capacity, 51, 132
Catchment analysis, 31, 80, 146, 293, 332, 337, 442, 493, 494, 513, 539, 628
Central Limit Theorem, 267
Central Place Theory, 20, 21, 75, 332, 339, 501, 515, 531, 532
Choice Theory (Rational Choice Theory) 41, 47, 48, 52
Christaller, 20, 21
Chi-square staristic, 66, 197, 218, 234, 241, 363, 400, 406, 608
City-block distance measure, 374
Clark, J. G. D., 29
Clarke, David L., 29

Class I overviews, 35, 41, 67, 69, 78, 80, 81, 85, 86, 287, 557, 617, 621, 630
Class II (sample) inventory, 43, 68, 78, 80, 270, 279, 628
Class III (intensive) inventory, 80, 446
Climate (see also Environmental factors), 23, 24, 30, 41, 159, 338, 453, 633, 635
Cluster analysis, 68, 156, 174, 208-210, 454, 456, 457, 538, 552, 621
  agglomerative hierarchical cluster, 208, 210
  distance measure, 209
  partitioning technique, 210
  Q-mode analysis, 208
  similarity coefficient, 209
Collectors (see also Foragers), 35, 114, 133-135, 137, 141, 289, 292, 321, 383
Collinearity, 215, 361
Colorado, 9, 26, 157, 195, 220, 290, 331, 338, 340, 356, 357, 363, 366, 378, 388, 411, 415, 416, 456, 518, 519, 521, 529, 531, 535, 610, 616, 624, 628, 632
Colorado Plateau, 26
Columbian Plateau, 26
Commerce, 33
Community patterning, 30
Comparative method, 24
Competitive exclusion principle, 141
Complex societies, 81, 258, 274, 283, 332
Compliance (see also Section 106; Section 110), 2, 3, 7, 8, 257, 550, 572, 575
Computer simulation, 82
Conditional probability, 180
Confidence
  interval, 44, 198, 214, 262, 316, 403, 404, 416, 560, 565, 632
  level(s) of, 6, 99, 531
  limits, 9, 16, 303, 316, 322, 400, 403, 404, 439
  map, 538
  mathematical vs explanatory, 99
Conklin, 29
Constancy (see also Resource distribution), 27, 131, 139, 141, 143, 552
Constrainrs on choice, 42, 79, 133, 339
Continental Shelf, 46-48, 581, 612
Contingency (see also Resource distribution), 28, 131, 139, 143, 400, 402, 552
Contrinuous probability distribution, 174, 196
Continuous variables, 183, 186, 333, 351
Contracts
  scopes of work, 282
Corps of Engineers (COE), 34, 65, 609, 613, 622
Correlated variables, 211, 222, 228, 388
Correlation, 4-6, 8, 9, 14, 24-26, 29, 33, 50, 75, 99, 100, 107, 109, 115, 122, 128, 132, 139, 157, 185, 202, 203, 208, 209, 211, 212, 214, 222, 225, 226, 229, 231, 232, 234, 238, 269, 278, 280, 358, 362, 388, 395, 439, 446, 447, 454, 459, 465, 482, 555, 559, 564, 567, 577, 616, 633, 637-639
Correlation matrix, 214, 225, 226
Correlative models, 4-6, 8, 9, 14, 15, 19, 35, 37, 52, 98, 99, 108, 127, 142, 160, 173, 264, 301, 325, 327, 330, 339, 386, 387, 415, 447, 451, 454, 465, 466, 478, 528, 556, 567, 573, 574, 576-578, 580, 613, 614, 624, 628, 630, 633
  limitations of, 33

Cost, 6, 16, 21, 34, 62, 97, 126, 151, 237, 262, 272, 286, 290, 292, 302, 303, 305, 317, 353, 372, 390, 395, 396, 406, 417, 433, 436, 439, 482, 494, 497, 499, 505, 506, 531, 538, 540, 553, 554, 557, 564, 568, 575-578, 580, 615, 616
Covariance matrix, 222, 225, 226, 234, 235, 237, 243, 246, 273, 367, 370
Crew spacing, 275, 304, 309
Critical resources, 32
Cross-tabulation, 202
Cultural ecology, 19, 23, 26, 28-29
  complexity, 29
  connectivity and mutual causality, 29
  monism, 29
  populations as analytic units, 29
  process, 29
  reorientation of, 29
Cultural resource management (CRM), 6, 7, 10, 32, 33, 37, 62, 66, 68, 69, 71-73, 75, 78, 81, 87, 88, 100, 122, 126, 144, 147, 160, 257, 259, 278-281, 286, 293, 294, 331, 469, 493, 497, 585, 609, 623, 632, 637, 638, 641
Cultural variables, 6, 7, 107, 122, 327, 439, 454, 456, 457, 463, 515, 528, 551, 584, 623
Culture and environment (see also Cultural ecology)
  relationship between, 25
Culture core, 26, 28
Curated and expedient technology, 118, 120, 121, 154

D
D'Holbach, 24
Damas, 25
Darsie, 30
Darwin, 24
Data
  aerial, 79, 430, 431
  accumulative, 280, 282
  accuracy, 287, 288
  activity area, 284
  analog,, 431
  archaeological, 24, 40, 65, 78, 79, 101, 102, 104, 126, 152, 154, 155, 159, 264, 270, 284, 304, 318, 339, 347, 350, 361, 364, 435, 438, 450, 451, 457, 463, 465, 468
  artifact, 147, 149, 150, 154, 155, 284, 291
  attribute, 155
  bias, 99, 257, 302, 304, 306, 308, 316, 394, 429
  chronological, 86
  clusters, 309, 396, 454
  comparable, 280, 287, 305
  control group, 346, 347, 349, 416
  cost, 290, 302, 497 cultural, 287, 289
  definition, 284
  density, 147, 153
  diagnostic, 289
  dimensions, 284
  display, 493, 494

Data
distributional, 154, 155, 158
ecomomic, 86
empirical, 301, 318, 320, 326, 330, 338, 341, 345, 347, 360, 386, 389, 400, 409
environmental, 70, 71, 78-80, 85, 104, 287, 301, 332, 415, 429, 437, 438, 447, 455, 493
errors, 291
ethnoarchaeological, 79, 104
ethnographic, 24, 78, 104
evaluation, 81, 86, 87, 313, 314, 317, 339
excavation, 39, 80, 81, 330
existing, 34, 41, 65, 76, 78-81, 83, 85, 86, 88, 257, 301-304, 306, 308, 314-317, 322, 383, 396, 432
feature, 284, 285
fit (see also Goodness-of-fit), 268, 341, 378, 381, 394, 400
flawed, 41
functional, 289
gaps, 86, 257
generalization of, 86
geographical, 382, 500, 501, 517-520, 529, 564
geomorphic, 86
hard, 34
hierarchical, 284
high-resolution, 155
historical, 78, 104
hyperfitting, 394
improvement, 307
inconsistencies, 301, 304
independent, 87, 257, 303, 314, 316, 392, 395, 399, 403, 407, 411, 437, 455
integration, 104, 316
interpretation, 101, 104, 121, 159, 386, 485
jackknife, 399
knowledge from, 98
limits, 290
locational, 150, 287, 301, 339, 350, 360, 372, 378, 399, 414, 481
macroenvironmental, 257
management, 292, 451
measurement, 126, 496, 497
microenvironmental, 257
MSS, 434, 440
new, 85, 87, 88, 257, 273, 290, 301, 302, 314, 316, 317, 355, 399, 400, 406, 417, 432
nonsite, 155, 287, 347, 360, 365, 372, 373, 381, 392, 411, 416, 417, 497
organization, 284
photographic, 430, 431
prevention of loss, 311
private, 306
problems, 304
processing, 290, 494, 518, 536
project-specific vs agency-wide, 87
proportions, 308

Data
quality of, 79, 87, 257, 287, 302, 304, 306, 307, 316, 390
quantitative, 339, 363, 386
quantity of, 79, 287, 302
raw, 386
reformating, 85
regarding functional dimensions of site classes, 86
regional, 79, 304, 318, 330, 331
relationships, 284
reliability, 151
remote-sensing, 38, 156, 157, 371, 429, 430
representative, 71, 85
requirements, 42, 293, 301
restructuring, 290
sample, 69, 71, 99, 147, 258, 261, 267, 270, 273, 313, 316, 317, 353, 366, 372, 374, 395, 399, 406, 407
satellite (see also Landsat, SPOT), 70, 430, 432-435, 476, 480
scale, 284, 288
secondary error, introduction of, 307
sequential or flat files, 292
sets for model testing and building, 43
settlement, 86, 329
simulated, 372
site, 41, 152, 284, 285, 291, 317, 345, 360, 365, 370, 372, 373, 381, 417, 481, 497
site content, 301
site location, 265
site-class definition, 289
social, 86, 332
sources of, 78, 79, 257, 466
spatial, 155, 352, 358, 493
spectral, 346
standardization, 288, 374, 387
static, 102
statistical, 292
storage, 292
subsampling, 308
subsistence, 80, 86
survey, 37, 39, 71, 81, 154, 257, 260, 282, 284, 287, 291, 301, 302, 304, 330, 384, 446, 450, 451, 457
temporal, 289
test, 71, 73, 83, 87, 126, 257, 303, 316, 358, 392, 395, 396, 399, 402, 403, 405, 411, 449, 497
topographic, 289
training, 350
transport, 292
unbiased, 98
unincorporated in regional plans, 257
useful for planning, 34
Data analysis, 31, 37, 65, 76, 98, 102, 122, 126, 145, 154, 159, 284, 307, 314, 315, 317, 325-327, 339, 345, 358, 360, 374, 381, 386, 394, 395, 429-431, 434, 444, 493, 494

Data collection, 46, 78, 85, 88, 97-99, 101, 121, 122, 126, 145, 147, 150, 151, 159, 257-259, 264, 265, 277, 280, 282, 284, 287, 288, 290, 292, 293, 303, 316, 317, 400, 406, 429-433, 449, 453, 493, 496, 497
  congruent with theory, 98, 99, 148, 281, 345
  cost-effective, 126, 151
  field, 284, 288, 292, 293
  for modeling, 72, 83, 86, 287, 316
  in CRM contexts, 277
  laboratory, 302
  multistep, 294
  on basis of intuition, 65
  on-site discovery rates, 308
Data recording, 65, 79, 122, 149-151, 282, 287, 288, 290-292
  code sheets, 291
  computerized, 46, 157
  consistency of, 287
  field recording, 291
  flexibility, 288
  hand-completed forms, 291
  inconsistency in, 151
  laboratory, 288, 292
  microcomputer, 292
  optical mark, 291
Data synthesis, 80
Dean, 32
Debitage, 154
Decision
  boundaries, 350, 351, 360, 361, 372, 377, 390, 416, 565
  region, 350, 360
  rule, 42, 314, 344, 350, 355, 360, 365, 366, 373, 374, 400, 403
  theory, 21, 40-43, 47, 48, 51, 72, 79, 80-82, 84, 85, 106, 107, 175, 220, 332, 335, 349, 582, 631
Deductive logic and/or models, 1, 2, 4, 14, 32, 37-39, 41, 46-48, 50-52, 63, 64, 67, 72, 73, 75, 76, 156, 158, 301, 303, 325, 386, 531, 554-556, 561, 564, 574, 576-578, 580, 582, 583, 612, 621, 622, 630
Density-contouring algorithms, 285
Dependent variables, 15, 63, 64, 69-72, 82-85, 203, 204, 213, 221, 243, 251, 326, 339-341, 371, 394, 437, 454, 463, 529, 550, 567, 611
Depositional processes (see also Archaeological record; Discard behavior), 37, 44, 79, 82, 102, 123-128, 143, 145-147, 152, 155, 157, 158, 259, 260, 275, 276, 281, 293, 434, 466, 467, 469, 475, 476, 574, 577, 584, 615, 639
Dichotomous response variable, 219, 340, 371
Differential equations, 61, 84
Diffusion, 25
Digital image processing, 158, 431, 435
Dimensionality, 185, 233
Discard behavior (see also Depositional processes), 102, 155, 466
Discrete variables, 183, 186

Discriminant
function analysis, 44, 45, 69, 73-75, 173, 174, 183, 210, 213, 220-222, 224, 225, 233, 235, 237, 242, 247, 248, 251, 320, 355, 364, 366, 371, 372, 382, 387, 394, 396, 415, 455, 456, 554, 575, 579, 616, 617, 628, 629, 632, 633
scores, 227, 229, 235, 237
Discriminating variables, 73, 221, 223, 224, 228, 233-235, 237, 247, 363
Distance to drainage variable, 346, 357, 364, 479, 497, 518, 613, 622, 624, 629, 633
Distance to vantage variable, 364, 624
Division of labor, 26
Dolores River Valley, 20
Domestication of plants and/or animals (see also Agriculture), 23, 134, 330
Double-blind test(s), 87, 287

E
Economic intensification (see Intensification)
Ecosystem, 25
definition of, 130
Ecosystemic
perspective, 29, 105
variables, 7, 107, 122, 128, 134, 143, 160, 434, 466, 478, 479, 482, 485, 552
Edge effect (see also Crew spacing; Sampling), 262, 264, 482
Eigenvalue, 227, 232
Elevation variable, 352, 479, 511, 516, 518, 622
Ellen, 25, 26, 29
Empirical/correlative models (see Correlative models)
Empiricism, 28
Environmental
determinism, 25, 28
factors, 5, 21-25, 27, 72, 82, 107, 135, 141, 142, 205, 258, 332, 356, 456, 519, 550, 559, 637-639
variables, 4-6, 10, 20, 28-30, 31, 33, 44, 66-68, 70-72, 74, 80, 81, 84, 107, 122, 127, 128, 132, 136, 138, 139, 142, 144, 200, 205, 208, 210, 220, 248, 279, 288, 318, 327, 330, 349, 357, 364, 377, 417, 435, 438, 439, 444, 446, 447, 455, 456, 459, 460, 463, 466, 479, 497, 515, 529, 531, 555, 567, 579, 584, 611, 620, 622, 628, 629, 632, 633, 637-639
Eskimo
Central, 25
Estimation
definition of, 68
least-squares method, 216, 217, 218, 219
maximum likelihood method, 195, 219, 364, 367, 374, 377, 390, 454, 455, 482, 538
parameter, 16, 44, 68, 69, 219, 281, 372, 560
Ethnoarchaeologists, 37, 112
Ethnographers, 37, 277
Ethnographic
analogy, 40, 79
context, 47, 79, 81, 104, 107, 108, 111-115, 143, 145, 156, 257, 345, 388, 466, 532, 608, 617, 618, 631, 632, 635, 639
models, 39, 608
record, 4, 41, 51, 608, 618, 639
research, 30, 75, 78, 79

Ethnohistory, 41, 78, 113, 145, 257, 621
Euclidean distance, 373
Evans-Pritchard, 23
Evolution,
  biological, 24
  cultural, 24
Evolutionary ecology, 29, 46
Executive Order 11593, 34
Explanation, 1, 3, 4, 6, 8, 9, 14, 22, 24-26, 29, 37, 61, 67, 75, 100-102, 104-106, 121, 128, 154, 158, 160, 161, 195, 212, 215, 220, 259, 386-389, 401, 429, 463, 465, 466, 479, 481, 555, 574, 577, 624, 633-635, 638
Explanatory
  models, 4, 6-9, 14, 37, 52, 61, 105, 265, 463, 573, 576, 578, 580
  variables, 213, 214, 216, 218, 219
Exploratory data analysis, 201-203
Exponential distribution, 196, 198
Exposure variable, 624

F

F-to-enter/F-to-remove, 238, 241
Factor analysis, 185, 208, 210, 212
Factorial, 191
Falsifiability, 15, 39, 102, 564, 583, 622, 632
Fisher's linear discriminant function, 234
Flannery, 29
Food resources, 32, 40, 42
Foragers (see also Collectors), 35, 50, 113, 130, 133-135, 137, 141, 142
Foraging radius, 109, 110, 112, 113, 118, 119, 134, 135
Forest Service, 9, 10, 34, 75, 79, 97, 535, 608, 618, 619, 621, 625, 626, 631
Formation processes (see Archaeological record; Transformation processes)
Fuel resources, 32, 40
Functionalism, 23

G

Gain
  definition of, 329, 344
  statistic, 329, 344, 347, 361, 365, 366, 371, 374, 377, 378, 381, 404, 565
Gamma distribution, 196, 199
Generalist encounter strategy, 110, 117, 133
Generalizability, 5, 6, 15, 39, 42, 50-52, 67, 583
Generalizing in science, 4-7, 9, 15, 16, 19, 23, 38, 39, 40, 42, 45, 47, 48, 50-52, 67, 69, 72, 74, 85, 86, 127, 174, 175, 181, 182, 189, 195, 200, 234, 241-243, 246-248, 250, 257, 258, 261, 265, 279, 281, 302, 386, 388, 435, 437-439, 447, 465, 466, 470, 476, 478, 563, 565, 574, 583, 612, 614, 618, 622
Genotype, 46
Geoarchaeology, 123, 293, 469
Geographic determinism, 23, 24
Geographic information system (GIS), 10, 17, 38, 46, 51, 97, 98, 144, 186, 302, 313, 318, 320, 349, 352, 363, 364, 366, 383, 384, 416, 439, 444, 451, 485, 493-543, 557, 568, 578, 579, 583, 610, 622-624, 638

Geographic information system (GIS)
analytic surfaces, 501
data planes, 501
digital elevation models, 502
raster or cell-based, 504
vector-based, 504
Geometric distribution, 191, 192
Geography, 20, 21, 24
Geological Survey (USGS), 333
Geomorphology, 260, 275
GIS (see Geographic information system)
Glacken, 22
Gold, 21
Goodness-of-fit (see also Statistical, fit), 66, 363, 378, 381, 405
Grass Mesa, 20
Gravity model, 75, 338, 531
Great Basin, 26, 27, 79, 82, 110, 113, 145, 284, 581, 585, 632-634
Great kiva, 338
Green,
Dee, 75, 499, 504, 625
Ernestine, 33, 332, 345, 387
Paul, 370
Stanton, 29
Gross error (see also Type I and Type II errors), 62, 63, 274, 347, 390, 565, 566
Ground steepness variable, 516

H
Habitation site, 283, 338, 519, 555, 562, 566
Haggett, 21
Hardesty, 29
Harris,
Marvin, 24, 28
Haury, 30
Heizer, 30
Hierarchical analysis, 41, 47, 52, 72, 382, 383
Higgs, 31
Hill,
James, 31, 32
Histograms, 200, 313
Historic Sites Act, 34
Historical particularism, 24, 25
Hodgen, 22, 23
Hohokam, 20, 80, 258, 274, 276
Human ecology, 22, 29
Human geography, 75
Human population
aggregates, 33, 130, 631
analytic units, 29, 346
change, 78, 555, 633, 639

Human population
density, 24, 26, 31, 112, 132, 133, 195, 475, 617, 618, 633, 639
distribution, 25, 26, 30, 31, 37, 131
genetic traits, 46
interaction, 32
packing, 136, 141
size vs site size, 48, 109
Hunter-gatherers, 5, 8, 15, 29, 35, 39, 44, 51, 106, 112-114, 130, 133-135, 146, 258, 284, 332, 335, 337, 345, 356, 387, 466, 479, 581, 611, 625, 631, 635
Hybrid variables, 185, 222
Hydrographic variables, 345
Hydrologic variables, 442
Hypergeometric distribution, 193

I
IDIMS (Digital Image Information System), 536
Inclusive fitness, 46
Independent testing, 303, 395, 416, 567
Independent variables, 5, 15, 33, 40, 63, 64, 67, 69, 71, 72, 81, 83, 85, 107, 120, 127, 139, 142, 173, 174, 181, 185, 202-204, 208, 210, 211, 213-216, 237, 243, 251, 387, 434, 437, 439, 463, 466, 478, 550, 561, 568
Inductive logic and/or models, 1, 4, 14, 28, 35, 37-39, 41, 46, 51, 52, 63, 64, 72, 73, 76, 142, 156, 264, 279, 447, 465, 470, 478, 545-556, 561, 564, 567, 574, 576, 577, 580, 582, 583, 611, 613, 614, 622, 624, 628
Industrial location, 21
Inferential logic and/or models, 19, 31, 32, 35, 37, 38, 42-44, 48, 51, 52, 137, 142, 143, 174, 351, 355, 536, 539, 629
Intensification, 5, 27, 40, 50, 131-136, 138, 139, 141, 142, 159
Interagency Archeological Services (IAS), 34, 612, 613
Intercorrelation, 361
Intermountain Antiquities Computer System (IMACS), 79
Internal consistency, 15, 39, 564, 583
Intersection of event sets, 180
Interval level (see Measurement, interval)
Intuitive logic and/or models, 4, 35, 63-65, 68, 76, 125, 156, 207, 222, 248, 350, 554, 561
Inventory (see also Class II and Class III inventory)
100 percent, 2, 34, 275
purposive selection, 260, 294, 308
Inyo Mountains, 27
Isaac
Glynn, 108, 145, 146
Isolated artifact/find/occurrences, 43, 44, 144, 151, 153, 282-284, 305, 382, 383, 585, 608, 616, 620, 621, 625, 632, 633

J
Jackknife test (see Statistical validation)
Joint frequency distribution, 202

K
Kalahari Desert, 110, 480
King, 34, 196, 267
Kohl, 28
Kolmogorov-Smirnov (KS) one sample test, 197, 358
Kriging, 69, 80, 340, 614, 615
Kroeber, 25, 26
Kurtosis, 189, 197, 198, 200

L
Landform variable, 345, 357
Landsat (see also Remote sensing), 44, 70, 71, 79, 355, 432-485 passim, 517, 540, 615, 620, 630, 632, 633, 638
Least-cost model, 21, 39
Least-squares method (see Estimation)
Level-slice classifier, 373
Limp, 41, 42, 47, 51, 52, 72, 104, 406, 407, 409
Linear additive model, 42, 42
Linear equation, 83, 456, 620
Linear regression
  assumptions, 185, 208, 213-215
  general, 174, 213, 214, 216, 539, 554
Lithic quarry, 283
Lithic scatter, 43, 74, 246, 283, 306, 340, 383, 388, 410, 519, 529, 531, 552, 629
Local relief variable, 364
Location
  definition of, 175
Location reuse, 48, 135
Locational behavior, 32
Locational bias, 306, 308
Locational variables, 78, 81, 338
Logical models, 63, 85
Logistic mobility, 111, 113, 121, 130, 134
Logistic regression, 69, 73-75, 173, 174, 183, 210, 213, 216, 219, 242, 247, 248, 251, 320, 340, 341, 355, 360, 361, 363, 364, 371, 372, 378, 381, 384, 387, 406, 415, 444, 457, 460, 519, 531, 554, 624, 630
Logistical radius, 110
Lognormal distribution, 196, 198
Low-density artifact scatter, 286, 289, 290

M
Magnet sites, 20, 80, 81, 258, 260, 273-275, 280, 281, 293, 561
  definition of, 258
Mahalanobis generalized distance measure, 209, 234, 241, 244
Maine, 46, 48, 611, 612
Mainframe computers, 290, 292
Management decisions, 67, 73, 74, 87, 549-551, 559, 563, 571, 575, 578, 627, 640
Management needs
  definition of, 559
  influence on survey design, 277

Map-based variables, 186
Map interpolation techniques, 69, 72, 80, 508, 510
Mappable variables, 248
Marginalism, 42
Marsh, 24
Mathematical models, 74, 85, 186, 393, 463
Maya Indians, 33
Maximum likelihood method (see Estimation)
Mean, statistical, 189, 194, 196-198, 200-202, 214, 220, 224, 225, 227, 229, 237, 244, 261, 267, 268, 271, 272, 279, 311, 313, 332, 351, 357, 358, 361, 364, 365, 370, 373, 374, 377, 394, 395, 403, 404, 454, 620
Means analysis, 285
Measure of dispersion, 189
Measure of location, 189
Measurement
  categorical, 35, 216, 221, 251, 332, 335, 341, 350, 351, 457, 505, 506
  interval, 35, 182-184, 186, 187, 197, 199, 208-210, 216, 221, 232, 251, 291, 332, 337, 346, 371, 554, 558, 584
  level of, 35, 41-43, 64, 182-185, 186, 202, 209, 221
  nominal, 35, 39, 183, 199, 202, 208, 216, 251, 291, 332, 333, 340, 345, 346, 371, 496, 505, 513, 554, 584
  ordinal, 35, 41, 42, 48, 183, 184, 199, 202, 216, 291, 335, 371, 554, 584
  ratio, 35, 41, 43, 182, 184, 199, 208-210, 216, 221, 251, 554
Measurement space
  definition of, 350
Median, statistical, 189, 200, 357
Mesolithic, 317, 318, 320-322, 331, 383-385, 532
Mid-Atlantic, 48, 612
Middle-range theory, 102, 117, 144, 155, 158
Mierendorf, 39-41, 581, 617, 618
Minimum distance classifier, 373
Mini-site, 146
Mobility patterns, 4, 109-111, 146
Mode, statistical, 189
Model building
  definition of process, quantitative predictive models, 204
  steps, 76
Model performance, 16, 242, 244, 247, 302, 303, 314-316, 322, 327, 344, 347, 353, 355, 357, 361-363, 366, 371, 372, 377, 389, 392, 393, 395, 396, 399, 400, 404, 406, 410, 411, 414, 415, 521, 532, 564
Model refinement, 6, 10, 15, 21, 38, 45, 78, 87, 152, 282, 317, 493, 532, 552, 555, 560, 561, 579, 580, 627, 628, 632
Model testing, 10, 15, 38, 43, 87, 316, 347, 406, 408, 497, 528, 531, 532, 553, 559-561, 564-566
  theoretical vs statistical, 104
Model validation, 72, 75, 87, 174, 242-244, 246, 247, 355, 386, 565
Model verification, 2, 7, 38, 242, 243, 246, 251, 288, 308, 386, 476, 494, 583
Model(s)
  definition of, 61
  hierarchical clustering, 279

Model(s)
  multistep, 250
  types of, 15, 63, 85, 173, 554, 556, 560
  utility, 5, 6, 10, 32, 52, 61, 72, 73, 160, 329, 347, 400, 402, 405, 447, 556, 561, 567, 575, 633
Modeling
  definition of, 101
  goals, 308
Monism, 29
Montesquieu, 23, 24
Morgan, 24
MOSS (Map Overlay and Statistical System), 504, 535
Moundville, 258
Multicollinearity, 185, 208, 214, 215, 221, 222
Multinomial distribution, 193
Multiple activity sets, 123
Multiple linear regression, 33, 42, 69, 71, 157, 210, 242, 371
Multistep sampling, 261, 294
Multivariate
  classification, 370, 394, 538
  distributions of variables, 201
  logistic models, 216
  mathematical techniques, 84, 325, 374, 403
  measure of group differences, 233, 395
  normal distribution, 183, 223, 234, 355, 371, 373
  normality assumption, 235, 364, 367
  parametric techniques, 208
  pattern-recognition model, 273
  predictive model(s), 43, 203, 212, 320, 322
  site location models, 78, 322
  statistical models, 33, 182
  statistics, 173, 174, 181, 182, 200, 202, 203, 208, 212, 224, 234, 238, 313, 314, 322, 327, 355, 358, 364, 370, 583, 609, 624, 638

N
National Environmental Policy Act (NEPA), 34
National Historic Preservation Act (NHPA), 7, 34, 550, 575, 634
National Park Service (NPS), 148, 149, 440, 442, 447, 450, 451, 470, 475, 482, 615, 627
National Register of Historic Places, 7, 34, 449, 550, 562, 627, 629, 630, 634, 638
Natural environment, 21, 23, 24, 26
Natural processes (see also Postdepositional processes; Taphonomic processes), 101, 123, 125, 126, 128, 151, 152, 159, 259, 275, 467, 469, 635
Natural selection, 29
Navajo, 114, 148, 152, 153, 470, 620
Negative binomial distribution, 194-196, 268
Netting, 29
New England, 48, 612
New Mexico, 70, 114, 122, 148, 344, 442, 447, 450, 451, 455, 462, 470, 620, 626
Neyman Type A distribution, 195
Neyman's allocation (see also Sampling), 272

No collection survey/policy, 16, 86, 151, 286
Nominal (see Measurement, nominal)
Nondimensional variable, 186
Nonsite
  distributional archaeology, 119, 122, 144-148, 155, 207, 283, 347, 478, 608
  site absent, 46, 74, 144, 174, 181, 204, 207, 208, 220-222, 225, 228, 229, 234, 235, 244, 248, 261, 283, 287, 314, 326, 329, 347, 350, 351, 355-358, 360-417 passim, 442, 446, 497-499, 528, 531, 616, 617, 624
Normal (Gaussian) distribution, 196-198, 200, 202, 222-224, 345, 355, 403
Normal equation, 218
Northern Paiute, 27, 111

O
Objective models, 64, 65
Objective(s)
  analytic, 174
  choice of, 76, 260, 555
  definition, 85
  management, 10, 68, 78, 83, 86, 88, 100, 257-259, 282, 286, 549, 551, 555, 567, 571-573, 609, 610, 622, 624, 626, 628, 634
  mapping, 35
  measurement, 80, 85
  model, 1, 6, 35, 63, 65, 554, 583
  of distributional archaeology, 148
  procedures, 206, 207, 210, 248, 350, 405, 494
  research, 7, 68, 76, 83, 86, 87, 100, 237, 257-259, 261, 264, 265, 277, 281, 282, 457, 549, 551, 555, 609, 610, 615-617, 619, 621, 622, 624, 626, 628, 632, 634
Objectivity, 28, 386
Occupation
  diversity, 206
  duration of, 48, 119, 133, 135, 444
  function of, 109, 124, 135, 143, 206, 289, 301, 321, 344
  intensive, 20, 48, 135, 206, 250, 446
  multiple, 33, 48, 115, 116, 124, 125, 283, 289, 444, 467, 475, 616, 626, 629, 632
  peak of, 20
  recurring, 114, 115, 259, 385
  season, period or periodicity of, 119, 123, 135, 301, 321, 467, 611, 617, 623
  single, 116, 123, 146, 157, 467
Off-site archaeology, 146, 148
Operationalization, 27, 32, 35, 39, 41, 42, 63, 64, 68, 76, 126, 143, 146, 207, 304, 313, 330, 333, 335, 382, 390, 430, 460, 532, 554, 555, 621
Optimal allocation, 273, 293
Optimal foraging theory, 42, 46-48, 50, 51, 612
Optimal location model, 75
Optimizing behavior, 31
Ordinal (see Measurement, ordinal)
Orthogonal variable, 204
Owens Valley, 27

P
Paleoenvironment, 47, 50, 86, 143, 557, 561, 584, 612, 614, 621
Paleoenvironmental reconstructions, 29
Paleoindian, 6, 276, 308, 330, 344, 475, 614, 616, 617, 620, 621
Parker, 19, 31, 39, 42, 81, 104, 138, 326, 335, 339-341, 371, 372, 387, 404, 405, 518, 610
Parsons, 19
Pascal distribution, 191, 192
Pastoralists, 112, 114
Patchiness (see also Resource distribution), 28, 131, 138, 552
Paynter, 20
Percent correct statistic, 403, 404, 411
Perlman, 47
Perspectives
  management, 19, 35, 62, 63, 67, 68, 72-74, 78, 83, 86-88, 101, 128, 160, 270, 275, 277-279, 331, 389, 609, 612, 616, 621, 629, 633, 634, 636, 640
  research, 19, 74, 78
Pixel, 72, 435, 485
  definition of, 349
Plant community variable, 351, 440, 629
Platform mounds, 274
Plog, 31, 32
Point-specific models, 63, 73-75, 261
Poisson distribution, 191-194, 196, 404
Polychotomous variables, 219
Polythetic choice model, 377
Polythetic-satisficer models, 75
Population (see Human population)
Population (see also Sampling; Statistics)
  of interest, 181, 261
Population genetics, 46
Postdepositional processes, 37, 79, 82, 85, 99, 102, 124-128, 143, 146, 147, 152, 155-157, 259, 260, 275, 281, 293, 429, 434, 466, 467, 469, 476, 478, 533, 557, 561
Precision
  concept of, 15, 39, 48, 51, 205, 554, 562, 564, 565
  degree of, 265
  gains in, 46, 326, 347
  geographic, 554
  level(s) of, 6, 553
  relative to specific model, 40, 41, 43, 45, 50, 51, 565
  statistical, 215, 242, 243, 262, 556
Predictability (see also Resource distribution), 21, 28, 100, 106, 109, 128, 131, 136, 138, 139, 142, 552
Prediction
  definition of, 2, 101, 104
Predictive locational modeling
  definition of, 33
Predictor variable, 142, 173, 182, 185, 204, 215, 247, 349, 361, 372, 394, 395, 406, 614
Principal components analysis, 44, 157, 185, 204, 208, 210-212, 222, 361, 362, 388, 444, 457, 463, 538

Private land, 273, 280, 305
Probability
  density function, 186, 187
  distribution model, 72
  mass function, 186, 187, 190-192
  theory, 176
Procedural logic, 35, 63, 64, 554, 555, 619
Projection, 5, 48, 109, 127, 248, 345, 416, 430, 433, 435-438, 441, 442, 444, 446, 447, 451, 457, 463, 499, 507, 516, 564, 610, 614, 632
Proton magnetometry, 126, 430
Pueblo, 6, 20, 114, 207, 219, 308, 338, 470, 609, 611

Q
Quadrat (see Sample Units, quadrat)
Quantitative
  method(s), 16, 156, 173, 175, 181, 201, 314, 337, 339, 346, 386, 497, 565, 574
  model(s), 16, 173, 175, 181, 204, 301, 320, 325, 327, 353, 363, 529, 532, 539, 556, 565, 574, 624, 637

R
Random
  experiment, 175, 176, 190, 191
  model, 41, 329
  variables, 173, 186, 187, 189-192, 198, 199, 213, 361, 394, 399, 403
Range of variability, 71, 242, 258
Rao's V, 241
Rappaport, 29
Ratio (see Measurement, ratio)
Rational Choice Theory (see Choice Theory)
Ratzel, 24
Reductionism, 26
Redundancy, 4, 123, 133, 479, 584, 634
Reed, 62, 66, 73, 261, 335, 416, 535
Reese River Valley, 27, 269, 284, 345, 377
Regional models, 26
Remote sensing, 10, 16, 17, 38, 70, 73, 79, 98, 126, 274, 313, 349, 355, 429-485, 494, 500, 516, 517, 532, 540, 557, 568, 578, 579, 615
  definition of, 430
  registration of data sources, 499
  scales and resolution, definition of, 431
Reoccupation, 115, 133, 475
Replicability, 386
Request for proposal (RFP), 150
Residential base, 50, 109-113, 115, 118-121, 130, 133-135, 137, 142
Resistivity measurement, 126
Resource distributions, 30, 122, 480
Response variable, 74, 207-209, 215-217, 219
Roberts, 35, 46-48, 581, 611
Roper, 31

S

Sample design, 87, 289, 327, 447, 451, 521, 528, 559, 566
cluster, 74, 261, 309, 351-353, 355, 358, 390, 394-396
element, 74, 261, 262
fraction, 86, 88, 205, 264, 267, 269, 270, 272, 274, 275, 289, 623
multistep, 250, 265, 277, 282, 293, 294
proportional, 624
purposive selection, 273
random, 44, 86, 182, 193, 197, 220, 259, 267, 269, 302, 307, 316, 317, 322, 327, 346, 353, 355, 356, 371, 389, 399, 400, 402, 409, 482, 529, 535, 558, 564, 566, 608, 609, 614, 616, 617, 623, 624
simple random, 86, 149, 207, 233, 259, 268-273, 276, 279, 293, 302, 309, 311, 351-353, 355, 361, 399, 528, 609, 629, 632
sample size, 16, 44, 86-88, 105, 193, 197-199, 203, 205-208, 223, 241, 243, 244, 246, 247, 261, 262, 264, 265, 267-269, 272, 273, 289, 290, 311, 316, 322, 329, 352, 364, 372, 382, 389, 393, 400, 404, 406, 409, 410, 416, 498, 499, 552, 566, 609, 629
stratified, 86, 145, 259, 269, 270, 273, 274, 276, 293, 447, 449, 608, 610, 614, 623, 624
subsampling, 276, 308, 312, 316, 395, 396, 528, 553
systematic, 248, 270, 528, 616
Sample survey (see also Class II inventory), 2, 67, 80, 86, 259, 260, 265, 268, 272-274, 276, 278, 293, 302, 393, 626, 628, 629, 633, 637, 640
Sample units, 73, 74, 86, 105, 147-151, 153, 158, 183, 216, 261, 262, 265, 267, 269, 270, 277, 283, 289, 290, 293, 305, 396, 409, 478, 552, 553, 558, 559, 564, 617, 620
quadrat, 33, 43-46, 86, 146, 184, 191-196, 208, 210, 247, 248, 262, 264, 265, 267, 268, 287, 308, 349, 390, 395, 609, 624, 632, 633
shape, 16, 86, 261, 262, 265, 289, 290
transect, 150-153, 259, 262, 264, 276, 290, 389, 457, 458, 481, 584, 617, 629
Sample universe, 265, 269, 270, 279, 289, 293
Sampling, 2, 5, 10, 16, 19, 33, 34, 37, 43, 68-70, 76, 80, 85, 86, 98, 122, 142, 144, 145, 147, 148, 173-175, 179, 197, 214, 242, 246, 247, 258-260, 265, 351, 356, 407, 409, 481, 556, 584, 609, 628, 640
as distinct from predictive modeling, 2, 68
bias, 305, 321, 388, 389, 566
cost, 2, 272, 303
diminishing returns of, 76
for independeut tests, 404, 410
pilot study, 272
probabilistic, 67, 71, 72, 74, 174, 182, 205, 233, 250, 260, 274, 278, 280, 289, 293, 294, 302, 303, 565
relationship to testable hypotheses, 2, 37
representative, 85, 302-304, 306, 314, 316, 318, 322, 327, 345, 346, 351, 352, 363, 384, 447
space, 180
splitting samples, 87, 174, 243, 315
theory, 260, 272, 303
variance, 215, 268, 271, 279, 370, 552
Sanders, 20
Santa Cruz River Valley, 20
SARG (see Southwestern Anthropological Research Group)
SAS (Statistical Analysis System), 370, 374
Satisficing models, 42, 48

Schiffer, 32, 33, 37, 82, 123, 259, 262, 264, 265, 280-282, 286, 291, 294, 304-306, 469
Schroedl, 43-45, 349, 364
Sears, 30
Section 106 (see also National Historic Preservation Act), 7, 8, 34, 550, 572, 575, 576
Section 110 (see also National Historic Preservation Act), 7, 8
Semple, 24
Settlement pattern studies, 19, 30, 31, 79, 205
Settlement system studies, 30
Settlement systems, 25
Settlements
  selective processes, 27
Shelter variable, 357, 364, 624
Shoshone, 27, 111, 113, 145, 344, 608
Significance
  archaeological, 7, 34, 125, 154, 159, 160, 466, 562, 565, 578, 584, 613, 617, 618, 622, 623, 628, 630, 634, 635, 638, 640
  archaeological, definition of, 331, 611, 630, 635
  statistical, 41, 197, 198, 206, 214, 223, 232, 233, 235, 238, 241, 303, 311, 316, 352, 355, 358, 389, 400, 402-404, 406, 410, 521, 608
Simple results, 175, 179
Simplicity, 15, 38-40, 43, 45, 50, 51, 67, 583, 616
Simulation model, 72, 608
Single-class classifier, 314, 345
Site location explanation
  agriculturalists, economic and cultural factors, 31
  causal vs culture-area, 26
  central place model, 20
  climate and environment, 23
  core features and secondary features, 26
  cultural factors, 19
  defensibility, 27
  distance costs, 21
  distribution of resources, 20
  economic, 21, 26, 27
  environmental, 21, 23, 28, 42, 137
  explanation after-the-fact, 29
  historical debate on, 22
  importance of nonfood resources, 32
  intensity of cultivation, 20
  interaction of variables, 29
  nonenvironmental, 20
  policy and education, 23
  political/governmental, 20, 23
  productivity of the soil, 20
  regional large-scale correlations, 26
  seasonally predictable, 28
  spacing of settlements, 20
  tailored, 27
  technology, 26

Site location explanation
  testing, 27, 31
  theological, 22
  topographical, 20
  transportation, 20, 21
  type of agriculture, 20
Site types, 28, 33, 39-41, 46, 50, 68, 82, 115, 122, 134-136, 139, 142, 205, 258, 260, 264, 269, 286, 293, 314, 316, 317, 321, 331, 345, 381-384, 417, 518, 528, 531, 541, 552, 557, 560, 563, 564, 569, 584, 585-611, 612, 616, 618, 620, 622, 625, 628, 630, 633-635, 637, 639, 640
Site visibility, 28, 86, 112, 116, 119, 121, 127, 133, 135, 138, 142, 143, 152, 184, 259, 276, 281, 289, 305, 306, 313, 321, 389, 429, 434, 463, 639
Site(s)
  classification, 68, 70, 71, 74, 81, 85-87, 174, 204-206, 208, 264, 265, 286, 314, 350, 370, 384, 385, 390, 403, 410-412, 414, 415, 531, 552, 560, 620
  clusters, 309
  concept of, 133, 138, 175, 144, 282, 284, 289, 301, 304, 326, 621
  discovery rates, 276, 304, 308, 313
  historical, 66, 205, 330, 441, 457, 621-623, 628, 630
  multiple occupations, 32
  organizational characteristics of, 32
  predictability, 28
  sample unit accessibility, 86, 273, 305, 306, 309, 563
  situation of, 31
  size, 43, 48, 109, 115, 119, 120, 175, 264, 281, 288, 584, 609, 612
  survivability, 275
Skewness, statistical, 44, 189, 197, 200, 206, 264, 265, 267, 275
Slope variable, 352, 358, 364, 440, 479, 497, 518, 613, 624, 629
Smith, 21
Snaketown, 258
Social factors, 80, 110, 327, 332, 345, 356, 566
Social variables, 21, 29, 205, 338, 356
Soil Conservation Service (SCS), 79, 433, 454, 476, 535
Soil-geomorphic model, 82
Soil variables, 345, 440, 442, 455, 518, 613, 622, 629
Southwest (see American Southwest)
Southwestern Anthropological Research Group (SARG), 31-33, 79, 338
Spatial referent, 15, 63
Spatial resolution (see also Remote sensing), 41, 47, 51, 432, 434, 436, 584, 612, 622, 632
Specificity, 10, 50, 83
Spencer, 24
Spiro, 258
SPOT (see also Remote sensing; Data, sattelite),
SPSS (Statistical Package for the Social Sciences), 224, 247, 311, 494, 539
SSCP/sums-of-squares-cross-products matrix, 210, 224-226, 235
Standard deviation, 189, 197, 198, 200, 201, 220, 227-229, 267, 269, 357, 377
State Historic Preservation Office(s), 11, 619
State Historic Preservation Officer(s), 34, 550
Statistical
  analysis of archaeological records, 175

Statistical
analysis packages, 244, 247, 273, 292, 311, 364, 382, 539
and environmental concept and measurement, 84
application success, 303
approaches to modeling, 16, 181
associations, 33
assumption of linear relationship, 85
assumptions about data, 181, 183, 214, 221, 273, 314, 351, 560, 566
classification procedures, 204, 233, 372
comparisons, 32, 41
compound processes, 195
confidence, 99
control-group approach, 314, 346
descriptive techniques, 174, 200, 251, 313, 317, 345, 386, 387
discussion, 14, 303
estimation procedure, 195
fit (see also Goodness-of-fit),
group identification, 74
independence, 67, 203, 352, 393
inductive procedures, 156
inferential techniques, 31, 174, 200, 247, 260, 303, 351, 352, 355, 536, 539
manipulation, 10, 15, 497
methods, 38, 76, 173, 325
models, 39, 75, 185, 203, 218, 251, 405, 528, 540
nonparametric techniques, 174, 199, 200, 203, 320, 364, 372, 554, 560
outliers, 45, 201, 202, 229, 373
parametric techniques, 174, 183, 199, 200, 203, 364, 554, 560
partial correlation, 203
performance, 327, 393, 400
probability, 174, 260
procedures, 71, 84, 181, 251, 313, 400, 402, 405
regularity of chance phenomenon, 176
relationship to scale of measurement, 184
sample size, 197, 198
scales of measurement review, 173
sophistication not leading to success, 75, 76, 181
stability of relative frequencies, 176
summary techniques, 386
technique, 10, 16, 17, 69, 85, 87, 638
terms discussed, 44, 62
test of fit, 218
theory, 84, 173, 303, 314
Statistical validation, 41, 242, 243, 246, 247, 355, 386, 393, 394, 531, 565
independent data, 87, 174, 242-244, 247, 250, 257, 392, 395, 403, 407, 553, 565
jackknife test, 220, 244, 315, 316, 395, 396, 399, 403, 414, 565
simulated data, 246
split-sample data, 174, 243, 316
Stepwise forward/backward selection procedure, 45, 237, 362, 363, 456
Steward, 15, 19, 26-28, 31, 79, 145, 585

Stewart, 26
Stochastic variable, 186
Storage, 39, 40, 109, 114, 120, 130, 133-135, 138, 139, 141, 292, 499, 504-506, 515, 535
Structure coefficients, 229, 231
Subsampling, 65, 244, 276, 308, 309, 312, 316, 317
Subsurface observations, 82, 124-126, 151, 159, 275, 276, 281, 294, 330, 467, 469, 554, 558, 568, 585, 613, 614, 623, 639
Sullivan, 32, 33
Superorganic, 24, 26
Surface
  paleo, 84, 260, 275, 276, 293
Surface observations, 82, 123-127, 145-148, 150-153, 159, 184, 205, 206, 259, 260, 274-276, 281, 286, 289, 330, 333, 337, 430-433, 440, 451, 453, 454, 467, 469, 470, 475, 476, 478, 480, 501, 505, 553, 554, 565, 566, 568, 581, 613, 614, 626, 634, 639, 640
Survey biases
  access difficulty, 305
  disproportionate areas, 312
  region underrepresented, 312
Survey intensity, 86, 277, 280-282, 293, 304, 305, 453
Survey universe (see also Sample universe), 16, 261, 264, 274, 277-279
  definition of, 278
Systemic context, 3, 8, 37, 38, 41-43, 47, 50, 51, 82, 577, 640
Systemic perspective, 30

T
t-distriburion, 196, 198, 214
t-test, 317, 318, 358
Taphonomic processes (see also Postdepositional processes), 115, 123, 146, 147
Target context, 35, 37, 64
Teorihuacán, 20, 258, 259, 274
Terrain variable, 349, 357
Testing (see also Model testing)
  null models, 32
  statisical comparisons, 32
Theory, 19-22, 24, 37-39, 41, 42, 46-48, 50-52, 64, 69, 72, 81, 84, 97, 98, 100, 102, 117, 128, 148, 156, 158, 260, 264, 272, 277, 280, 286, 303, 314, 332, 338, 349, 355, 367, 386, 404, 612
Thomas, 27
Tipps, 43, 44, 71, 265, 279, 581, 632, 633
Topographic variable, 351, 442, 622
Trade, 27, 33, 132, 134, 135, 137, 566, 630, 634
Transect (see Sample units, transect)
Transformation processes (see also Postdepositional processes)
  natural and cultural, 638
Trend-surface analysis/mapping, 80, 181, 330, 339-341, 344, 378, 538
Trigger, 28, 30, 31
Turgot, 24
Tylor, 24
Type I and Type II errors
  definition of, 62

U
Ullman, 20
Uncorrelated variables, 204, 211, 212
Underground radar, 126
Univariate statistical techniques, 174, 181, 200, 313, 320, 609, 622
Use intensity, 4
USGS (see Geological Survey)
Utah, 43, 79, 265, 267, 278, 416, 444, 456, 609, 628, 632

V
Validation (see Model validation; Statistical validation)
Variable(s) (see also Cultural, Environmental, Dependent, etc.)
  correlations among, 185
  definition of, 183
  selection of, 16, 185
  spatial distribution of, 69
Variance, statistical, 74, 155, 157 182, 189, 194, 196, 200, 202, 210-212, 214-216, 218, 222, 225, 237, 241, 246, 262, 268, 270-273, 279, 311, 313, 333, 355, 358, 360, 364, 370, 372, 402, 442, 454, 456, 480, 513, 552, 613, 620
Vegetation (see Ecosystemic variables; Environmental variables; Plant community variables)
  Verification (see Model verification)
VICAR (Video Image Communication and Retrieval), 537
View variable, 357, 358
Virú Valley, 30
Vita-Finzi, 31
Von Thünen, 20

W
Washington (state), 39, 40, 609, 617, 618, 631
Wasteful error (see also Type I and Type II errors), 62, 347, 390, 565, 566
Weber, Alfred, 21
Weibull distribution, 196, 199
Weighted
  analysis, 65, 308, 311, 316
  least-squares estimation, 218
  sample mean, 311
  variance, 311
White, 26
White Mountains, 27
Wildesen, 34, 582, 634-636
Wilks's lambda, 233, 241
Willey, 30
Williams, 30
Winterhalder, 29, 47
Wissler, 25
Woodland, 613

Z
z-scores, 352

☆ U.S. GOVERNMENT PRINTING OFFICE:1988-576-901/85245