New Approaches to the Use and Integration of Multi-Sensor Remote Sensing for Historic Resource Identification and Evaluation

SERDP Project SI-1263

University of Arkansas, Center for Advanced Spatial Technologies

Kenneth Kvamme, Eileen Ernenwein, Michael Hargrave, Thomas Sever, Deborah Harmon, Fred Limp

Lead Principal Investigator Fred Limp

November 10, 2006

Distribution Statement A: Approved for Public Release, Distribution is Unlimited

This report was prepared under contract to the Department of Defense Strategic Environmental Research and Development Program (SERDP). The publication of this report does not indicate endorsement by the Department of Defense, nor should the contents be construed as reflecting the official policy or position of the Department of Defense. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the Department of Defense.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to the Department of Defense, Executive Services and Communications Directorate (0704-0188). Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.
PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ORGANIZATION.

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
15-11-2006	Final	1/2002- 11/2006

4. TITLE AND SUBTITLE
New Approaches to the Use and Integration of
Multi-Sensor Remote Sensing for Historic Resource Identification and Evaluation

5a. CONTRACT NUMBER
DACA72-02-C-0035

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Kenneth Kvamme, Eileen Ernenwein, Michael Hargrave, Thomas Sever, Deborah Harmon, Fredrick Limp, Burgess Howell, Michele Koons, Jason Tullis

5d. PROJECT NUMBER
CS-1263

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of Arkansas, Center for Advanced Spatial Technologies
12 Ozark Hall
Fayetteville, AR 72701

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Strategic Environmental Research and Development Program
1155 Herndon Parkway, Suite 900
Herndon, VA 20170

10. SPONSOR/MONITOR'S ACRONYM(S)
SERDP

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited

13. SUPPLEMENTARY NOTES
Prepared in cooperation with the U of A Department of Anthropology, ERDC CERL and NASA Marshall Space Flight Center

14. ABSTRACT
This is the final report on a SERDP research project whose focus is the identification of specific combinations of remote sensors and data integration methods for the detection, identification, and interpretation of cultural resources in various environments and archaeological circumstances. The extensive suite of sensor technologies employed for this task include magnetometry, magnetic susceptibility, electrical resistivity, electromagnetic conductivity, ground penetrating radar, aerial thermal infrared and high resolution multispectral satellite imagery. The results of the research illustrate the very substantive subsurface site characteristics that are discoverable from integrative methods used. Based on these results a dramatically clearer picture of the subsurface is realized, compared to traditional site evaluative methods. By more clearly imaging the totality of information about the subsurface from all sources, a better understanding of site content, structure, and organization may also be achieved.

15. SUBJECT TERMS
Archaeology, Remote Sensing, Geophysics, Data Fusion, Cultural Resource Management, Magnetometry, Magnetic Susceptibility, Electrical Resistivity, Conductivity, GPR, Thermal Infrared, Satellite Imagery

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE		392	19b. TELEPHONE NUMBER *(Include area code)*

Reset

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18

TABLE OF CONTENTS

ACKNOWLEDGEMENTS xvi
LIST OF ACRONYMS xviii
1. EXECUTIVE SUMMARY 1
1.1 GENERAL SUMMARY 1
1.2 PROBLEMS ADDRESSED 1
1.3 SCIENTIFIC QUESTIONS 2
1.4 RESULTS AND FUTURE APPLICATIONS 3
2. OBJECTIVE 6
2.1. CULTURAL RESOURCE MANAGEMENT DETECTION & EVALUATION TECHNOLOGIES: THE SERDP STATEMENT OF NEED 6
2.2. REMOTE SENSING FOR ARCHAEOLOGICAL DETECTION AND EVALUATION 7
2.2.1. Remote Sensing and Archaeology 7
2.2.2. Pattern Recognition and Large-area Surveys 8
2.2.3. Benefits of Remote Sensing 9
2.2.4. The Nature of Archaeological Sites 10
2.2.5. Properties of Archaeological Sites: What Can Be Remotely Detected?. 12
2.2.6. Sensor Mixes and Detection Probabilities 13
2.2.7. Prognosis 13
2.3. PROPOSED RESEARCH: INTEGRATED ARCHAEOLOGICAL REMOTE SENSING FROM LAND, AIR, AND SPACE 14
2.3.1. Data Types Investigated 15
2.3.2. Benefits of Multiple Detection Methods and Data Integration 16
2.3.3. Project Objectives 18
2.3.4. Deliverable Products 20
3. BACKGROUND 22
3.1. DoD REGULATORY REQUIREMENTS AND PROJECT FOCUS 22
3.1.1. DoD Regulatory Requirements 22
3.1.2. Project Problem Focus and Regulatory Requirements 23
3.1.2.1. Deficiencies of current archaeological discovery methods 24
3.1.2.2. This Project's benefits and regulatory requirements 25
3.2. HISTORY OF DATA INTEGRATION IN ARCHAEOLOGICAL REMOTE SENSING 26
3.3. SELECTING PROJECT SITES 27
3.3.1. Site Selection Issues 27
3.3.2. Site Selection Goals 29
3.3.3. Candidate Sites 30
3.3.4. Selected Project Sites 31
3.3.4.1. Army City, Kansas 31
3.3.4.2 Pueblo Escondido, New Mexico 36
3.3.4.3. Silver Bluff Plantation, South Carolina (The George Galphin Site) 39
3.3.4.4. Kasita Town, Georgia 41
4. MATERIALS AND METHODS 45
4.1. GEOPHYSICAL THEORY 45

4.1.1. Active and Passive Methods 45
4.1.2. Anomalies 45
4.1.3. Area Surveys and Sampling Density 46
4.1.4. Magnetometry 47
4.1.5. Electrical Resistivity 48
4.1.6. Electromagnetic Induction 49
4.1.7. Ground-penetrating Radar 50
4.2. GEOPHYSICAL FIELD SURVEY METHODS 51
4.2.1. Survey Blocks 52
4.2.2. Data Acquisition and Sampling Density 52
4.2.3. Survey Issues 53
4.3. GEOPHYSICAL INSTRUMENTATION 55
4.3.1. Magnetometry 55
4.3.2. Electrical Resistivity 56
4.3.3 Electromagnetic Induction 58
4.3.4. Ground-penetrating Radar 59
4.4. GEOPHYSICAL DATA PROCESSING 60
4.4.1. Noise and Defects 61
4.4.1.1. Noise inherent to a site 62
4.4.1.2. Instrument noise 63
4.4.1.3. Operator errors 66
4.4.2. Enhancements 68
4.4.2.1. Interpolation 68
4.4.2.2. Low-pass filtering 69
4.4.2.3. Contrast manipulation 70
4.4.2.4. Shadowing 71
4.4.3. Presentation 71
4.4.3.1. Continuous gray or color scales 72
4.4.3.2. Other modes of portrayal 73
4.4.4. Processing GPR Data 74
4.4.4.1. Processing GPR profiles 75
4.4.4.2. GPR depth determination 76
4.4.4.3. GPR time- and depth-slices 78
4.4.5. Processing Sequences for Geophysical Data 79
4.4.5.1. GPR processing sequence 79
4.4.5.2. Processing electrical resistivity data 80
4.4.5.3. Processing EM conductivity data 82
4.4.5.4. Processing EM in-phase (magnetic susceptibility) data 83
4.4.5.5. Processing magnetic gradiometry data 84
4.5. FORMING INTERPRETATIONS FROM REMOTE SENSING DATA 86
4.5.1. Prior Assumptions 86
4.5.2. Anomaly Identification through Excavation 86
4.5.3. Anomaly Identification through Pattern Recognition 87
4.5.3.1. Pattern recognition based on shape 87
4.5.3.2. Pattern recognition by systematic repetition 88
4.5.3.3. Pattern recognition by relative size 89

4.5.3.4. Pattern recognition by association 90
4.5.3.5. Pattern recognition by context 91
4.5.3.6. Experience as a basis for pattern recognition 91
4.5.3.7. Pattern recognition approaches: summary 93
4.5.4. Anomaly Identification through Theory 93
4.5.4.1. Natural causes of environmental magnetism 94
4.5.4.2. Anthropogenic causes of magnetic variation 94
4.5.5. Discussion 98
4.6. GEOPHYSICAL RESULTS 98
4.6.1. Geophysical Results at Army City 98
4.6.1.1. Anomalies common to all data sets 100
4.6.1.2. Electrical resistivity 101
4.6.1.3. GPR 101
4.6.1.4. EM conductivity 101
4.6.1.5. Magnetic gradiometry 102
4.6.1.6. Magnetic susceptibility 102
4.6.2. Geophysical Results at Pueblo Escondido 102
4.6.2.1. Magnetic Methods 104
4.6.2.2. Electrical Methods 105
4.6.2.3. Ground-penetrating Radar 106
4.6.2.4. Advanced computer-visualization of the Escondido GPR data 107
4.6.3. Geophysical Results at Silver Bluff Plantation 111
4.6.3.1. Magnetic gradiometry 112
4.6.3.2. Electrical resistivity 113
4.6.3.3. Ground-penetrating radar 114
4.6.4. Geophysical Results at Kasita Town 114
4.6.4.1. Magnetic gradiometry 115
4.6.4.2. Electrical Resistivity 116
4.6.4.3. Ground-penetrating Radar 116
4.7. AIR AND SPACE REMOTE SENSING THEORY & HISTORY 118
4.7.1. Theory 118
4.7.1.1. Thermal infrared 119
4.7.2. Archaeological Aerial Photography 120
4.7.2.1. History 120
4.7.2.2. Archaeological air photo interpretation 121
4.7.3. Aerial and Space Imaging Activities in the SERDP Project 124
4.8. AERIAL & SPACE REMOTE SENSING BY UNIVERSITY OF ARKANSAS 124
4.8.1. Powered Parachute & Instrumentation 125
4.8.2. Orthorectification 126
4.8.3. The Potential of Aerial Remote Sensing at Army City 126
4.8.4. Normal-light Aerial Photographs and Imagery from Army City 131
4.8.5. Thermal Infrared Imagery from Army City 132
4.8.6. Thermal Infrared Results from Pueblo Escondido 134
4.8.7. Relationship between Satellite Panchromatic, Aerial Thermal, Conductivity, and GPR data at Pueblo Escondido 135

4.9. AERIAL AND SPACE REMOTE SENSING BY NASA 138
4.9.1. Introduction 138
4.9.2. Sensor Instrumentation 138
4.9.2.1. QuickBird 138
4.9.2.2. Raytheon PalmIR 250 138
4.9.2.3. DuncanTech MS4100 Digital Multispectral Camera 139
4.9.2.4. Advanced Thermal and Land Applications Sensor (ATLAS) 139
4.9.3. Image Analysis Software 140
4.9.4. Processing Techniques 140
4.9.4.1. Georeferencing 140
4.9.4.2. Resolution enhancement 141
4.9.4.3. Image stacking 141
4.9.4.4.Band ratioing 141
4.9.4.5. Filtering 141
4.9.4.6. Principal Components Analysis (ATLAS thermal) 141
4.9.5. Data Analysis 142
4.9.6. Features of Interest 143
4.9.6.1. Army City 143
4.9.6.2. Pueblo Escondido 145
4.9.6.3. Silver Bluff 147
4.9.6.4. Kasita Town 150
4.9.7. Lessons Learned 153
4.9.7.1. Data acquisition 153
4.9.7.2. Time of acquisition of thermal data 153
4.9.7.3. Mode of acquisition of thermal data 154
4.9.7.4. Site preparation 154
5. RESULTS AND ACCOMPLISHMENTS 155
5.1. PROJECT HISTORY & MILESTONES 155
5.1.1 Objectives 155
5.1.2 Objectives Summary 156
5.1.3 Deliverable Products 159
5.2. PREPARING DATA FOR INTEGRATION 160
5.2.1. Site Selection and Methods 160
5.2.2. Pre-processing for Data Integration 161
5.3. TRADITIONAL INTERPRETIVE APPROACHES FOR DATA INTEGRATION 163
5.4. GRAPHICAL METHODS FOR DATA INTEGRATION 166
5.4.1. Graphic Overlays 166
5.4.2. RGB Color Composites 167
5.4.3. Translucent Overlays 168
5.4.4. Evaluation 170
5.5. DISCRETE INTEGRATING METHODS 170
5.5.1. Pre-processing: Generating Binary Data 170
5.5.2. Boolean Union 172
5.5.3. Boolean Intersection 173
5.5.4. Binary Sum 174

5.5.5. Thresholds on Binary Sum 174
5.5.6. Polychotomous Integrations 174
5.5.7. Cluster Analysis Methods 175
5.5.8. Evaluation 178
5.6. CONTINUOUS INTEGRATING METHODS: SIMPLE MATHEMATICAL APPROACHES 179
5.6.1. Data Sum 179
5.6.2. Data Product 180
5.6.3. Data Ratios 180
5.6.4. Data Maximum 181
5.6.5. Data Variance 182
5.6.6 Evaluation 182
5.7. CONTINUOUS INTEGRATING METHODS: MULTIVARIATE TECHNIQUES 183
5.7.1. Principal Components Analysis 184
5.7.2. Factor Analysis 187
5.7.3. Supervised Statistical Classifiers 189
5.7.3.1. Binary logistic regression 190
5.7.3.2. Multinomial logistic regression 192
5.7.6. Evaluation 197
5.8. INTELLIGENT KNOWLEDGE-BASED SYSTEMS FOR DATA INTEGRATION 198
5.8.1. Artificial Intelligence and Geocomputational Tools 198
5.8.2. C5.0 and Machine Learning Decision Trees 199
5.8.3. C5.0 Decision Tree Classification of Army City 202
5.8.3.1. C5.0 classification of standardized versus un-standardized data ... 202
5.8.3.2. C5.0 results at Army City 203
5.8.3.3. Winnowing, cross-validation, and boosting 205
5.8.4. C5.0 and Image Segmentation 206
5.8.5. Decision Tree Regressions using Cubist 208
5.9 AN EXPERT SYSTEMS APPLICATION 210
5.9.1 Expert Rules for Army City 211
5.9.2. Discussion and Evaluation 214
5.10. EVALUATION AND BENEFITS OF DATA INTEGRATION 215
5.11. DATA INTEGRATIONS FROM OTHER STUDY SITES 218
5.11.1. Pueblo Escondido 218
5.11.1.1. Integrating GPR data 219
5.11.1.2. Integration by summing ranked polychotomies 220
5.11.1.3. Integration by binary summation 221
5.11.1.4. Integration by the maximum function 222
5.11.1.5. Integration by statistical algorithms in supervised classification 222
5.11.2. Silver Bluff Plantation, South Carolina 224
5.11.2.1. Integration by RGB color composites 225
5.11.2.2. Continuous Data Integrations 226
5.11.2.3. Cluster analysis 227
5.11.2.4. Binary logistic regression 228

5.11.3. Kasita Town, Georgia 228
5.11.3.1. Binary data integrations 230
5.11.3.2. Continuous data integrations..... 230
5.11.3.3. Binary logistic regression 231
5.11.3.4. K-means cluster analysis..... 232
5.11.3.5. Conclusions..... 233
5.12. INTERPRETATIONS AND VECTORIZATION 234
5.12.1. Overview of Process 234
5.12.2. Methods..... 235
5.12.3. Evolving Perspectives..... 236
5.12.4. Army City 236
5.12.5. Pueblo Escondido..... 240
5.12.6. Silver Bluff..... 243
5.12.7. Kasita Town 245
5.13. ISSUES IN FIELD VALIDATION OF ANOMALY PREDICTIONS 247
5.13.1. Issues in Anomaly Source Validation and "Ground Truthing" 248
5.13.2. Field Assessment Tactics..... 251
5.13.3. Research Plan..... 252
5.14. SAMPLING DESIGNS FOR ARCHAEOLOGICAL FIELDWORK..... 253
5.14.1. Types of Excavation Samples..... 255
5.14.2 Sampling Army City 256
5.14.2.1. Types and numbers of samples (excavations) 256
5.14.2.2. Sampling design..... 257
5.14.3. Sampling Design for Pueblo Escondido 259
5.14.4. Sampling Design for Silver Bluff 262
5.14.5. Sampling Design for Kasita Town..... 264
5.15. VALIDATION PHASE FIELD METHODS 266
5.15.1. A Multi-Staged Strategy for Anomaly Validation..... 267
5.15.2. The SERDP Project's Field Validation Strategy 268
5.15.3. Field Preparation and Fieldwork..... 270
5.16. ARCHAEOLOGICAL INVESTIGATION OF REMOTE SENSING PREDICTIONS..... 273
5.16.1. Field Investigations at Army City..... 274
5.16.1.1. Evaluation of predictions by archaeological team..... 276
5.16.1.2. Evaluation of predictions by SERDP team..... 280
5.16.2. Field Investigations at Pueblo Escondido..... 292
5.16.2.2. Evaluation of predictions by SERDP team..... 301
5.16.3. Field Investigations at Silver Bluff..... 314
5.16.3.1. Evaluation of predictions by archaeological team..... 316
5.16.3.2. Evaluation of predictions by SERDP team..... 319
5.16.4. Field Investigations at Kasita Town 329
5.16.4.1. Evaluation of predictions by archaeological team..... 333
5.16.4.2. Evaluation of predictions by the SERDP team..... 334
5.16.5. Summary and Discussion..... 336
5.17. REVISED SITE INTERPRETATIONS..... 340
5.17.1. Revised Interpretations at Army City 340

5.17.2. Revised Interpretations at Pueblo Escondido 342
5.17.3. Revised Interpretations at Silver Bluff Plantation 345
5.17.4. Revised Interpretations at Kasita Town 346
6. CONCLUSIONS 349
6.1. BENEFITS OF REMOTE SENSING IN ARCHAEOLOGY 349
6.2. BENEFITS OF DATA FUSION 350
6.3. TOWARDS AN EPISTEMOLOGY OF ARCHAEO-GEOPHYSICAL FIELD VALIDATION 351
6.4. FIELD VALIDATION OR "GROUND-TRUTHING" 354
6.5. GROUND-TRUTHING VS. VALIDATION 355
6.6. DATA REDUNDANCIES AND COST-BENEFITS 355
6.7. EFFECTS OF ENVIRONMENT ON DETECTION 356
6.7.1. Urban Environments 356
6.7.2. Non-urban Settings 357
6.7.3. Aerial (and Space) Methods 357
6.7.4. Geophysics 357
6.8. BEST SUBSET OF INSTRUMENTS 358
6.8.1. Environment 358
6.8.2. Nature of Archaeological Remains 358
6.8.3. Considerations of Dimensionality 360
6.9 COST/BENEFIT OF TECHNOLOGY 360
6.10 TRANSITION STATUS 361
7. REFERENCES CITED 362
APPENDIX A: SUPPORTING DATA 371
APPENDIX B: LIST OF TECHNICAL PUBLICATIONS 372
Articles Published In Peer-Reviewed Journals 372
Technical Reports 372
Conference/Symposium Proceedings 372
Published Technical Abstracts 373
Published Book Chapters 374

LIST OF FIGURES AND TABLES

Figure 2.1. Regular geometric shapes in anomaly forms point to cultural constructions in this ground-penetrating radar image from Pueblo Escondido..... 9
Figure 2.2. Components of the archaeological record. 11
Figure 2.3. Prehistoric and historic archaeological sites revealed by remote sensing..... 12
Figure 2.4. Different geophysical methods yield complementary information.. .. 16
Figure 3.1. Present-day Army City showing portions of its former street grid and the study area centered over the town's commercial district. 33
Figure 3.2. Panoramic photograph of the commercial core of Army City shortly after its construction in 1917..... 33
Figure 3.3. Electrical resistivity survey of Army City..... 35
Figure 3.4. The site of Pueblo Escondido, New Mexico. 36
Figure 3.5. The Silver Bluff site. a) Aerial view showing loci of archaeological excavations and geophysical survey boundary. 39
Figure 3.6. Satellite view of the SERDP Project study area immediately west of Lawson airfield and above the Chattahoochee River 42
Figure 4.1. Two-dimensional mapping of an electrical resistivity data set from Army City, Kansas..... 46
Figure 4.2. Increasing sampling densities. 47
Figure 4.3. GPR field methodology and sampling 51
Figure 4.4. Geophysical survey methods. 54
Figure 4.5. Effects of ferrous metal debris and constructions on magnetic gradiometry data. 55
Figure 4.6. Magnetic gradiometry survey at Pueblo Escondido..... 56
Figure 4.7. Electrical resistance meters. 57
Figure 4.8. Electromagnetic induction survey at Army City 58
Figure 4.9. The GSSI SIR-2000 ground-penetrating radar system..... 60
Figure 4.10. Removal of data spikes..... 62
Figure 4.11. Historic or recent plow marks 63
Figure 4.12. Correction of apparent discontinuities..... 64
Figure 4.13. Correction of heading and striping errors..... 65
Figure 4.14. Staggering caused by operator timing errors..... 67
Figure 4.15. Effects of operator gait on magnetometry data sets.. 68
Figure 4.16. Interpolation improves image continuity, visualization, and the recognition of culturally-generated anomalies..... 69
Figure 4.17. Benefits of low-pass filtering.. 70
Figure 4.18. Contrast enhancement of magnetic gradiometry data from Pueblo Escondido..... 71
Figure 4.19. Relief shadowing can enhance image details. 72
Figure 4.20. Gray and several color scale displays of resistivity data..... 73
Figure 4.21. Typical modes of graphic display..... 74
Figure 4.22. Typical processing sequence for two radargrams 76
Figure 4.23. Determining GPR wave velocity..... 77

Figure 4.24. *GPR time-slices.* 79
Figure 4.25. Processing of Army City GPR data. 80
Figure 4.25.1. Processing of Army City electrical resistivity data. 81
Figure 4.26. Processing of Army City electromagnetic conductivity data. 83
Figure 4.27. Processing of Army City magnetic susceptibility data. 84
Figure 4.28. Processing of Army City magnetic gradiometry data. 85
Figure 4.29. Cultural features exhibit regular geometric shapes. 88
Figure 4.30. Cultural features may exhibit repetitive patterns. 89
Figure 4.31. Relative size differences associated with functionality. 90
Figure 4.32. Pattern recognition by association at small map scales 91
Figure 4.33. Interpreting anomalies based on prior archaeological experience. .. 92
Figure 4.34. The fired brick foundation of the Mount Comfort Church (1840s-1863), Arkansas 95
Figure 4.35. Accumulations of magnetically enriched topsoil generate positive anomalies. 96
Figure 4.36. The removal of magnetically enriched topsoil can produce negative anomalies. a 97
Figure 4.37. Interpretations of Army City geophysical data 100
Figure 4.38. Geophysical survey results at Pueblo Escondido. 105
Figure 4.39. Geophysical survey results at Pueblo Escondido. 106
Figure 4.40. GPR survey results at Pueblo Escondido. 107
Figure 4.41. Thirteen one-nanosecond thick time-slices, each representing approximately 16 cm of subsurface "thickness." 108
Figure 4.42. 3-D visualization with Slicer Dicer software.. 109
Figure 4.43. 3-D visualization of isosurfaces describing GPR anomalies with 3-D Master software. 110
Figure 4.44. 3-D visualization with MATLAB 7.0 111
Figure 4.45. Geophysical survey results at Silver Bluff Plantation. 113
Figure 4.46. Geophysical data collected at Kasita Town 116
Figure 4.47. Kasita Town ground-penetrating radar depth slices. 117
Figure 4.48. Phenomena that cause features of archaeological sites to be visible from the air. 122
Figure 4.49. Vegetation and shadow marks point to subsurface archaeological conditions. 123
Figure 4.50. Two-seat powered parachute 125
Figure 4.51. GCP and image rectification and registration. 126
Figure 4.52. Rectangular cultural feature of Army City 128
Figure 4.53. Stilted and browned grasses forming negative vegetation marks. . 128
Figure 4.54. Negative vegetation marks revealing linear architectural features. 129
Figure 4.55. Building footings revealed as negative vegetation markings at Army City. 130
Figure 4.56. Map of surface-visible features revealed by vegetation markings . 130
Figure 4.57. Aerial imagery from Army City.. 132
Figure 4.58. Thermal infrared imagery from Army City showing details of building footings. 133

Figure 4.59. Thermal infrared imagery from Army City showing wall and footing details of the Hippodrome. 133
Figure 4.60. Thermal infrared composite for Army City..... 134
Figure 4.61. Thermal infrared imagery over principal study block at Pueblo Escondido..... 135
Figure 4.62. Comparison between satellite and aerial data over the principal Pueblo Escondido study area. 136
Figure 4.63. Pueblo Escondido geophysical data and patterns of vegetation. 137
Figure 4.64. Full scene, true color QuickBird image, band 3,2,1 (RGB) showing the general area of study at Army City 143
Figure 4.65. Composite Army City imagery 144
Figure 4.66. Army City imagery. 145
Figure 4.67. Full scene, true color QuickBird image of the Pueblo Escondido. 146
Figure 4.68. Enhanced imagery of the Pueblo Escondido study region.. 147
Figure 4.69. Full scene, true color QuickBird image, bands 3, 2, 1 148
Figure 4.70. 149
Figure 4.72. Duncan multispectral data. 149
Figure 4.73. Full scene, true color QuickBird image, band 3,2,1 150
Figure 4.74. Pan sharpened and contrast enhanced QuickBird image..... 151
Figure 4.75. Pan sharpened and contrast enhanced QuickBird image. 152
Figure 4.76. Composite imagery at Kasita Town.. 153
Figure 5.1. Normalizing transform and standardization of the Army City resistivity data.. 162
Figure 5.2. The six pre-processed geophysical data sets from Army City: 163
Figure 5.3. Interpretive vectors overlaid on geophysical data and each other.... 165
Figure 5.4. Graphic overlays as data integration. 167
Figure 5.5. RGB color compositing.. 168
Figure 5.6. Translucent overlays..... 169
Figure 5.7. Flattening/binary classification of the Army City resistivity data.... 171
Figure 5.8. Binary representations of principal geophysical anomalies at Army City..... 172
Figure 5.9. Discrete data integrations at Army City. 173
Figure 5.10. Integration of ranked polychotomous geophysical data by summation.. 175
Figure 5.11. K-means cluster analyses of the Army City data 178
Figure 5.12. Continuous data integrations. 180
Figure 5.13. *Data ratios.* 181
Figure 5.14. a) Largest anomalies in the six geophysical data sets obtained through application of the maximum function..... 182
Figure 5.15. Mapped results of principal component analyses..... 186
Figure 5.16. Factor analysis of the Army City data. 189
Figure 5.17. Performance of the binary logistic regression model. 192
Figure 5.18. Modeling specific archaeological classes with multinomial logistic regression for data fusion..... 194
Figure 5.19. C5.0 rules generated for the extraction of residential dwellings 199
Figure 5.20. C5.0 used for extracting (classifying) impervious surfaces. 200

Figure 5.21. Hypotheses, rules, and conditions form the basic structure of production rules 201
Figure 5.22. Example of rules and the rule structure for a C5.0 classification.. 204
Figure 5.23. C5.0 classification output at Army City 205
Figure 5.24. Original scale and seven scales of multi-resolution image segmentation of Army City's geophysical data. 207
Figure 5.25. Classification of image segments 208
Figure 5.26. Example of a Cubist rule developed for predicting the sum of robust anomalies. 209
Figure 5.27. Comparison between the actual sum of all robust binary anomalies and the cubist-modeled estimates. 210
Figure 5.28. Expert model specifying geophysical rules and conditions for identification of 5 hypotheses. 212
Figure 5.29. Expert system model applied to Army City geophysical data. 213
Figure 5.30. Three-dimensional rendering methods applied to integrated data. 215
Figure 5.31. Fusion of the three GPR time-slices from Pueblo Escondido in the first principal component of a PCA analysis. 219
Figure 5.32. Ranked categories based on standard deviation thresholds..... 220
Figure 5.33. Data fusions based on discrete categorical methods at Pueblo Escondido..... 221
Figure 5.34. Data integrations at Pueblo Escondido based on the GIS maximum function. 222
Figure 5.35. Integration results at Pueblo Escondido by statistical algorithms based on supervised classifications..... 224
Figure 5.36. RGB fusion results at Silver Bluff Plantation. 226
Figure 5.37. Continuous fusion results at Silver Bluff Plantation. 227
Figure 5.38. Unsupervised and supervised classification results at Silver Bluff Plantation. 228
Figure 5.39. Integration of GPR depth slices by principal components analysis at Kasita Town. 229
Figure 5.40. Binary data integrations at Kasita Town. 231
Figure 5.41. Continuous data integrations at Kasita Town..... 231
Figure 5.42. Binary logistic regression results at Kasita Town. 232
Figure 5.43. K-means cluster analysis results for five clusters at Kasita Town. 233
Figure 5.44. Illustration of the process of vectorizing anomalies at Army City. 238
Figure 5.45. Vector interpretation for Pueblo Escondido. 243
Figure 5.46. Mapping of classified anomaly types for Silver Bluff Plantation. . 245
Figure 5.47. Vector interpretation for Kasita Town. 247
Figure 5.49. Anomalies illustrating clear and unambiguous archaeological features indicated by geophysics and corresponding archaeological manifestations. 249
Figure 5.50. Evidence for anomaly identifications at Army City..... 252
Figure 5.51. Archaeological sampling design and distribution of excavation types at Army City 259
Figure 5.52. Distribution of excavation units at Pueblo Escondido overlaid on the vector interpretation. 262

Figure 5.53 Distribution of excavation units at Silver Bluff overlaid on the vector interpretation……264
Figure 5.54. Distribution of excavation units for Kasita Town overlaid on the vector interpretation. 266
Figure 5.55. Excavation methods for field validation in the SERDP Project..... 268
Figure 5.56. Forms used in field evaluations at Army City..... 269
Figure 5.57. Validation phase fieldwork at Pueblo Escondido..... 271
Figure 5.58. Excavation unit form used by SERDP team to evaluate archaeological sources of remote sensing anomalies at Army City. 272
Figure 5.59. View of project excavations at the site of Army City 275
Figure 5.60. Trench 12 at Army City excavated to investigate a robust concentration of debris or rubble. 277
Figure 5.61. Concrete footer identified in a test unit at Army City. 278
Figure 5.62. Samples of archaeological sources for anomalies classified as "chimney base" and "footings." 281
Figure 5.63. Samples of archaeological sources for anomalies classified as "isolated metal" and "pipes." 282
Figure 5.64. Samples of archaeological sources for anomalies classified as "walls." 284
Figure 5.65. Samples of archaeological sources for anomalies classified as "floors" and "debris piles." 285
Figure 5.66. Samples of archaeological sources for anomalies classified as street and alley "gutters." 287
Figure 5.67. Samples of archaeological sources for anomalies classified as "street surface," "street rut," and "sidewalk edge." 288
Figure 5.68. Final distribution of units for Pueblo Escondido test excavations overlaid on the vector interpretation. 294
Figure 5.69. SERDP team members using a KT-9 Kappameter hand-held magnetic susceptibility meter 296
Figure 5.70. TRC crew uses a 3-inch auger to investigate anomalies at Pueblo Escondido. 299
Figure 5.71. The sharp edge of a pit house with relatively rich fill in Trench 2 at Pueblo Escondido. 300
Figure 5.72. Testing of a floor anomaly with mechanized trench M16. 303
Figure 5.73. Testing of a wall anomaly with mechanized trench M2. 304
Figure 5.74. Testing of an interior room feature anomaly with mechanized trench M13. 305
Figure 5.75. Testing of linear anomaly with mechanized trench M26. 306
Figure 5.76. Testing of two small feature anomaly with mechanized trenches M25 and M19.. 307
Figure 5.77. An example of distinct house anomaly that was nearly invisible in the test excavation. 312
Figure 5.78. An example of a distinct house anomaly for which no evidence was found by excavation. 313
Figure 5.79. Excavation of a backhoe trench at Silver Bluff..... 314
Figure 5.80. Profile map of Mechanized Trench 19. 317

Figure 5.81. Features revealed during Silver Bluff excavations..... 318
Figure 5.82. Archaeological testing of the "iron or steel" anomaly category..... 319
Figure 5.83. Evidence for the anomaly type "magnetic point anomalies in linear arrangement."..... 320
Figure 5.84. Area anomalies and corresponding evidence at Silver Bluff. 321
Figure 5.85. Positive resistivity anomalies and archaeological sources. 322
Figure 5.86. Linear GPR anomalies and their archaeological sources……323
Figure 5.87. Examples of post holes at Silver Bluff and associated anomalies.. 324
Figure 5.88. Evidence of a paleochannel 325
Figure 5.89. Resistivity anomalies of apparent anthropogenic origin compared to historic plow furrows. 328
Figure 5.90. Excavations at Kasita Town. 329
Figure 5.91. Photograph of a portion of backhoe trench 2 showing the plow zone and spoil from military grading of the airfield..... 331
Figure 5.92. Map of cut and fill areas as a result of grading at Kasita Town..... 332
Figure 5.93. Photographs illustrating the main causes of anomalies at Kasita Town. 334
Figure 5.94. Revised classification of Army City anomalies based on knowledge gained in the archaeological testing program. 341
Figure 5.95. Reinterpretation of Pueblo Escondido based on testing of geophysical anomalies. 343
Figure 5.96. Revised classification of Silver Bluff anomalies based on knowledge gained in the archaeological testing program. 345
Figure 5.97. Reinterpretation of Kasita Town based on testing of geophysical anomalies. 346

Table 2.1. Basic Data for Instruments and Sensors Employed in Project. 16
Table 2.2. Principal SERDP Project tasks accomplished by season and year..... 19
Table 3.1. Characteristics of sites selected for inclusion in the SERDP study..... 31
Table 4.1. Geophysical methods, instruments, sampling densities, and principal anomaly types at Army City. 99
Table 4.2. Geophysical methods, instruments, sampling densities, and principal anomaly types at Pueblo Escondido. 103
Table 4.3. Geophysical methods, instruments, sampling densities, and principal anomaly types at Silver Bluff. 112
Table 4.4. Geophysical methods, instruments, sampling densities, and principal anomaly types at Kasita Town..... 115
Table 4.5. QuickBird bands and band widths. 139
Table 4.6. DuncanTech MS4100 bands and band widths..... 139
Table 4.7. ATLAS sensor system specifications. 140
Table 5.1 Subtask Schedule and Final Report Cross-Reference..... 157
Table 5.2. K-means cluster analysis statistics (n=64,000)..... 177
Table 5.3. Pearsonian correlation matrix between the six normalized geophysical variables. 184
Table 5.4. Principal component analysis of the standardized geophysical data. 185
Table 5.5. Factor analysis of the geophysical data, with varimax rotation..... 188

Table 5.6. Confusion matrices showing performance of the binary logistic regression model 191
Table 5.7. Parameter estimates for the four logistic regression functions 195
Table 5.8. Classification accuracies of the Army City models. 197
Table 5.9. Accuracy matrices showing C5.0 classification performance. 203
Table 5.10. Classification accuracy of expert model 214
Table 5.11. Pearson correlation coefficients between integrating methods..... 217
Table 5.12. Correlation matrix between geophysics acquired at Escondido. 219
Table 5.13. Correlation matrix between continuous fusions at Escondido..... 224
Table 5.14. Correlation matrix between geophysical data at Silver Bluff. 225
Table 5.15. Correlation matrix between geophysical data at Kasita Town. 229
Table 5.16. Description of 17 defined anomaly classes at Army City. 239
Table 5.17. Description of 15 defined anomaly classes at Pueblo Escondido.... 242
Table 5.18. Thirteen anomaly types at Silver Bluff Plantation..... 244
Table 5.19. Anomaly classes for Kasita Town vector interpretation..... 246
Table 5.20. Anomaly types, strata, population and sample sizes for the Army City excavation program..... 258
Table 5.21. Anomaly types, strata, population and sample sizes for the Pueblo Escondido excavation program..... 261
Table 5.22. Thirteen anomaly types at Silver Bluff Plantation..... 263
Table 5.23. Anomaly types, strata, population and sample sizes for the Kasita Town excavation program. 265
Table 5.24. Overall percentage of correct predictions for all excavation units. . 276
Table 5.25. Percentage rate of correct predictions for shovel tests and hand-excavation units 277
Table 5.26. Percentage rates of correct predictions for backhoe trenches..... 278
Table 5.27. Incidence of positive and negative results in efforts to verify the presence of features associated with each anomaly type. 279
Table 5.28. SERDP team evaluation statistics for Army City's anomaly types..289
Table 5.29. Comparison of accuracy rates between the limited exposures of the STP and the backhoe trench..... 290
Table 5.30. Distinctly defined archaeological features with sharp boundaries are more likely to be confidently assessed..... 291
Table 5.31. Anomaly types by the principal geophysical device. 292
Table 5.32. Actual number and distribution of units excavated at Pueblo Escondido..... 295
Table 5.33. Results of ground truthing for backhoe trenches at Escondido. 297
Table 5.34. Results of ground truthing for auger tests at Pueblo Escondido..... 297
Table 5.35. Results of ground truthing for backhoe trenches at Escondido. 298
Table 5.36. Summary of test excavation success at Pueblo Escondido..... 308
Table 5.37. Summary of test excavation success by geophysical method at Pueblo Escondido..... 309
Table 5.38. Summary of test excavation success according to the number of different geophysical indications. 310
Table 5.39. Summary of test excavation success by type of excavation used at Pueblo Escondido..... 310

Table 5.40. Cross tabulation of anomaly categories and feature types at Silver Bluff. 316
Table 5.41. Distribution of iron and steel artifacts in STP. 319
Table 5.42. Accuracy assessment of Silver Bluff anomalies by SERDP team... 326
Table 5.43. Accuracy assessment of geophysical survey types at Silver Bluff.. 327
Table 5.44. Summary of Kasita Town geophysical anomalies and excavation units. 330
Table 5.45. Summary statistics for Kasita Town anomaly testing. 335
Table 6.1. Characteristics of four principal subsurface prospecting methods. ... 358

ACKNOWLEDGEMENTS

The SERDP project team extends its sincere thanks to the many individuals listed below (and any who were inadvertently omitted) who contributed to the successful outcome of this research.

Strategic Environmental Research and Development Program
For the funding and technical guidance in undertaking this project
Mr. Bradley Smith, Executive Director
Dr. Robert Holst, Program Manager, Sustainable Infrastructure
HGL Support Staff – Ms. Susan Walsh, Ms. Kate Kerr, Mr. John Thigpen

Archaeo-Physics LLC, Minneapolis
David Maki, Geoff Jones

Powered Parachute at Army City
Tommy Ike Hailey, Cultural Resource Office, Northwestern State University, Louisiana

Fieldwork at Army City
Jennifer R. Bales, Eileen G. Ernenwein, Charles K. Kvamme, Jo Ann Kvamme, Dorothy Neeley, Christopher Goodmaster, Christine J. Markussen, Jack Cothren, Malcolm Williamson, Mohammed Salem, Debbie Harmon, Fred Limp, Robert Harris, Burgess Howell.
Dan Lee, GeoTek (pilot for NASA aerial imagery acquisition, Army City, Kasita Town and Silver Bluff)

Fieldwork at Pueblo Escondido
Eileen G. Ernenwein, Elsa Heckman, Jason Herrmann, Terri Bruce, Aaron Gunn, Alicia Valentino, Sara Gale, Christopher Goodmaster, Christine Markussen

Fort Benning
Dr. Christopher Hamilton, Cultural Resources Management (CRM) Program Director
Mr. John Brent, Chief, Environmental Management Division
Mr. Peter Drake, CRM program
Dr. Thomas H. Foster II, Ph.D., BHE Environmental Project Director
Dr. Christopher Bergman, BHE Environmental CRM Program Director
Mr. Mike Bruening, BHE Environmental graphics
Mr. Dean Wood, Southern Research Inc.
Mr. Matthew Wood, Southern Research Inc.

Bureau of Land Management
Ms. Pamela Smith, Archaeologist

Fort Bliss
Mr. Keith Landreth, Chief, Conservation Division
Ms. Vicki Hamilton, Division Chief, Conservation Division
Mr. Brian Knight, Senior Archaeologist
Ms. Sue Sitton, CRM program
Mr. Shane Offutt, CRM program
Dr. Bret Ruby, CRM program (now Cultural Resources Program Manager, Fort Lewis)
Mr. James Bowman, CRM program (now Cultural Resources Manager, White Sands)
Mr. Paul Lukowski, TRC Environmental-El Paso co-Project Director
Ms. Elia Perez, TRC Environmental-El Paso co-Project Director
Mr. Paul Webb, TRC Environmental-Raleigh, NC
Cherokee Enterprises, Alamogordo NM

Fort Riley
Mr. Herb Abel, Directorate of Environment and Safety (DES)
Mr. Scott Hall, CRM Program Director
Dr. Richard Shields, DES
Mrs. Fiona Price, CRM program
Mr. Craig Phillips, DES
Mr. John Dendy, CRM program
Dr. Paul Kresja, University of Illinois, PSAP Project Director

Savannah River Archaeological Research Program
Mrs. Tammy Forehand Herron, SRARP co-Project Director
Mr. Robert Moon, SRARP co-Project Director
Dr. Adam Smith, SRARP archaeologist
Mr. Tommy Charles, SCIAA

Silver Bluff Audubon Center and Sanctuary
Mr. Robert Connelly, SBACS Manager

Fort Drum
Dr. Laurie Rush, CRM Program Director

Fort Hood
Dr. Cheryl Huckerby, CRM Program Director

Geoscan Research USA
Dr. Lewis Somers (provided the hand-held probe array for testing at Kasita Town)

LIST OF ACRONYMS

AI	Artificial Intelligence
CERL	Construction Engineers Research Laboratory
CFR	Code of Federal Regulations
CRM	Cultural Resource Management
CON	Conductivity
DES	Directorate of Environment and Safety
DoD	Department of Defense
DoE	Department of Energy
EM	Electromagnetic (Induction)
EPAS	El Paso Archaeological Society
ESTCP	Environmental Strategic Technology Certification Program
GIS	Geographic Information System
GPR	Ground Penetrating Radar
HDM	Horizontal Dipole Mode
MAG	Magnetic Gradiometry
MS	Magnetic Susceptibility
NHPA	National Historic Preservation Act
NRHP	National Register of Historic Places
PCA	Principal Components Analysis
RDP	Relative Dielectric Permittivity
RES	Electrical Resistivity
RGB	Red-Green-Blue
SHPO	State Historic Preservation Office
SBACS	Silver Bluff Audubon Center and Sanctuary
SCIAA	South Carolina Institute of Archaeology and Anthropology
SERDP	Strategic Environmental Research and Development Program
SI	Systeme Internationale
SON	Statement of Need
SOW	Statement of Work
SRARP	Savannah River Archaeological Research Program
SSE	Sum of Squared Error
STP	Shovel Test Pits
THPO	Tribal Historic Preservation Officer
THERM	Thermal Infrared
TRC	TRC Environmental Inc.
TWTT	Two-Way Travel Time
USC	United States Code
VDM	Vertical Dipole Mode

1. EXECUTIVE SUMMARY

1.1 GENERAL SUMMARY

The Center for Advanced Spatial Technologies and the Department of Anthropology, University of Arkansas, ERDC CERL and NASA Marshall Center have conducted research in a project titled *New Approaches to the Use and Integration of Multi-Sensor Remote Sensing for Historic Resource Identification and Evaluation (SERDP CS-1263)*. The focus of this research is the identification of specific combinations of remote sensors and data integration methods for the detection, identification, and interpretation of cultural resources in various environments and archaeological circumstances. Four sites were selected for inclusion in this project: Army City (Fort Riley, Kansas), Pueblo Escondido (Fort Bliss, New Mexico), Kasita Town (Fort Benning, Georgia), and Silver Bluff (located near the DOE Savannah River facility, Aiken County, South Carolina). An extensive suite of ground-based geophysical, aerial, and satellite technologies were employed for this task. The methods investigated include magnetometry, magnetic susceptibility, electrical resistivity, electromagnetic conductivity, ground penetrating radar, aerial thermal infrared and high resolution multispectral satellite imagery. The goals of this research include:

- Application of a suite of established and newly fielded high-resolution space and aerial instrumentation to the detection of cultural resources.
- Examination of a suite of existing and newly developed ground based instruments applied to the detection and evaluation of subsurface deposits.
- Application of existing and the development of new computational approaches for the integration of an extensive suite of satellite, aerial, and ground-based remote sensing data sets.
- Assessment of the results of these approaches through an archaeological field validation program in a range of environments and archaeological site types.
- Evaluation of the effectiveness of various sensor combinations and fusion methods against site conditions, and development of guidance documents for their use in various settings.

1.2 PROBLEMS ADDRESSED

Some of the nation's most significant historic and prehistoric cultural resources are contained within the 25 million acres of public land administered by the Department of Defense. Protecting these heritage resources is a fundamental part of the Department's primary mission. Heritage management issues central to that mission have focused on the economics of identifying and maintaining historic facilities, the impact of archaeological sites on construction and training programs, and the disposition and curation of artifacts. Cultural resources, defined in DoD Instruction 4715.3, include buildings, structures, sites, districts, and objects eligible for, or included in, the National Register of Historic Places regulation (36 CFR 60). Management of these resources, in compliance with existing laws and regulations such as *Native American Graves Protection and Repatriation Act* (25 USC §3001), the *Archaeological Resources Protection Act* (16 USC §470 aa-ll), *American Indian Religious Freedom Act* (42 USC §1996), and the standards in the *Curation of Federally-Owned and Administered Archaeological Collections* (36 CFR

79), necessitates the development of innovative and cost-effective methods for archaeological site identification, evaluation, and protection.

The Strategic Environmental Research Development Program (SERDP) Statement of Need entitled "Cultural Resources Management Detection and Evaluation Technologies" (CSSON-02-01) requested research into a range of advanced archaeological methods. Specifically, it called for methods (1) that "effectively detect, locate and identify historic and pre-historic archeological resources on military and DoE lands and ranges," (2) that produce "improved models for predicting the location of resources," and (3) that identify "improved technologies for detecting surface and/or subsurface resources." Non-invasive processes and procedures were "strongly encouraged in order to reduce the possible disturbance of human remains and associated artifacts." Finally, validation of findings through excavation was deemed "necessary to demonstrate the feasibility of the proposed technologies and associated procedures."

This Statement was timely, because mainstream archaeological methods for the identification and evaluation of most historic resources remain little changed from those employed in the early twentieth century. Surface survey and excavation, the traditional field methods for discovery of artifacts, architectural elements, and other archaeological features, continue to predominate in spite of the fact that these techniques are extremely time consuming, expensive, and unreliable. Their shortcomings manifest themselves in many ways. Frequently there is no discernable surface evidence of buried archaeological features, making surveys ineffective. Excavation of small shovel test-pits can sometimes locate archaeological sites, but this method entails substantial additional survey costs. In National Register of Historic Places eligibility assessments, ever-increasing labor costs restrict the number of units that can be excavated, resulting in failures to locate features of significance. More importantly, traditional invasive methods regularly lead to the damage or destruction of the very resources they were designed to investigate and generate costs associated with collection curation. Aware of these limitations, archaeologists have adopted a conservative, preservationist approach, with the result that the Department of Defense protects many sites of marginal scientific importance. It is apparent that the SERDP SON implicitly recommends *archaeological remote sensing* as a primary vehicle for meeting current needs.

1.3 SCIENTIFIC QUESTIONS

A principal goal of this SERDP research project is a determination of remote sensing methods and techniques that work well individually, and that complement each other collectively when integrated, for the identification and evaluation of buried archaeological remains. Numerous methods for data integration were explored. With the exception of certain multi-band visualization techniques and overlays of vectors representing interpreted anomalies, most of the methods investigated have not been applied previously in archaeology. Several advanced computer graphic methods were explored, discrete methods that range from *Boolean* overlays to sums of categorized portrayals of sensor outputs and cluster analyses, continuous methods that include sums, products, ratios, principal components, regressions, and probability surfaces, intelligent knowledge-based systems like *C5.0* and *Cubist*, and Expert Systems approaches. This research demonstrates that certain integrating methods yield more information about the subsurface than others, but what may be realized in each approach may depend on

overarching purpose. Some fusion techniques yield visually pleasing results that appear to well-combine available information, for example, while others may seem less revealing but offer greater interpretive or predictive potential. In this process, the nature of similarities, differences, redundancies, and performance characteristics of results were examined. An important aspect of this research is an assessment of the added value of the fused product compared to traditional, individual-sensor based analysis.

1.4 RESULTS AND FUTURE APPLICATIONS

The process of archaeological remote sensing as carried out in this project is a multi-step undertaking. The first stage, designed to meet the primary data integration goal, includes remote sensing data collection, processing to clarify anomalies in individual sensors established within GIS databases, data fusion to integrate information from all sensors, definition of potentially "significant" cultural anomalies, and classification of the anomalies into likely types of cultural features. This is normally the final product of most remote sensing projects, and the point at which archaeological fieldwork takes over. This project endeavored to go several steps farther with a *second* principal goal.

The second goal of this research project was very different and designed to meet an important criterion in the SERDP Statement of Need. Specifically, "ground-truth" testing was called for to demonstrate the feasibility of the data integration technologies and associated procedures. Three additional project tasks were therefore designed. They include development of a sampling design that allowed archaeological excavation of representative anomalies of each defined type to provide validation of remote sensing predictions about the subsurface. This validation phase often turned into a learning process, however, because the soils, geology and the archaeology in each site are unique, idiosyncratic, and confound predictability. In other words, remote sensing predictions cannot be perfect and a look into the ground through excavation offers additional insights that allow modifications to original predictions. Consequently, a final stage was defined that includes modification of original remote sensing predictions, based on excavation findings. Numerous scholarly presentations and publications were produced during the project. Preparation of a final report was, of course, the ultimate task.

These operations were undertaken at each of four prehistoric and historic archaeological sites distributed across time and space in a wide diversity of environmental settings from South Carolina to New Mexico. This provided a variety of contexts in which to assess the value of the methods investigated at very different archaeological sites with very different remains in very different environments. In so doing a better understanding of which methods consistently worked and offered useful results could be achieved, but this knowledge was also augmented by the considerable experience of project team members.

Integrating multiple geophysical data sets offers large potential for improved understanding of the subsurface. A single survey, for example, might reveal only part of a buried building. Integrated information from several surveys may illustrate the entire structure as well as interior components. Moreover, integrated data may *simultaneously* show relationships between conductive, resistant, magnetic, thermal, and metallic anomalies, potentially improving knowledge of features within a site, inter-sensor relationships, enhancing overall interpretations.

Graphical solutions for data integration are easy to implement and effectively combine information from disparate sources into interpretable displays. They allow complex visualizations of the subsurface, but their weakness rests in relatively low dimensionality—only 2-3 data sources may effectively be represented. Moreover, these methods are purely descriptive, yielding only images, not new data that may subsequently be analyzed. Discrete integrating methods, on the other hand, allow application of readily available Boolean operations to any number of geophysical data sets. A shortcoming is that the binary maps upon which these methods are based rely on arbitrary thresholds to define significant anomalies, while more subtle ones must be ignored. Continuous data integrations can yield insights beyond the capabilities of other methods. Robust *and* subtle anomalies may be simultaneously expressed, producing composite imagery with high information content. Interpretive data are also generated in the form of principal component scores, factor loadings, or regression weights that add to understanding of interrelationships and underlying dimensionality. Supervised and unsupervised classification methods are noteworthy because they introduce a predictive aspect to the integrating process. Patterns in these data fusions may point to anomalous conditions much less visible in any single data set that might otherwise be overlooked. They therefore offer a possible means to augment prospecting capabilities. Although the approaches to geophysical data integration examined here span a wide range of commonly available techniques, they are by no means exhaustive. A host of other supervised and unsupervised classification algorithms exists, as well as new context-based image segmentation, and intelligent knowledge-based methods.

If the foregoing results can be generalized, it is that robust anomalies exist in the data and tend to dominate any form of fusion, regardless of the method employed. The consequence is amazingly parallel results between widely different forms of integration. Consequently, they really should be considered as offering new information about subsurface variation.

The determination of which integrating methods are best may depend on purpose. Some yield visually pleasing results that appear to well-combine available information while others may seem less revealing but may offer interpretive or predictive potential. If a goal is to define discrete classes of anomalies that may be subsequently interpreted through comparison with primary data then categorical methods may be best. If a goal is merely a continuous-tone image that represents most of what is known about the subsurface then a composite color graphic or mathematical-statistical integration may be most suitable. Of course, continuous methods yield quantitative data that may subsequently be analyzed, plus regression weights, PCA scores, or factor loadings give additional insights beyond graphical representations, and are important for improved understanding of the subsurface and its interaction with geophysical methods. In practice, a variety of different integrating methods may work best in practice, because each variation may give new insights about a different aspect of the subsurface.

The results of the integrated data sets clearly illustrate the very substantive subsurface site characteristics that are discoverable from the integrative methods used. Based on these results a dramatically clearer picture of the subsurface is realized, compared to traditional site evaluative methods. By more clearly imaging the totality of information about the subsurface from all sources, a better understanding of site content, structure, and organization may also be achieved. The amount of information provided

from these methods dramatically improves the ability to assess the site properties consistent with eligibility evaluations. The extensive amount of information yielded by the approach also will serve as important guidance should site mitigation be needed. Compared to the typical site evaluation results these methods provide orders of magnitude more information on the nature of the internal sites structures and its characteristics. This is particularly evident at sites such as Army City and Pueblo Escondido but to a lesser degree at all the sites.

It should be noted that while these methods are very effective in the horizontal delineation of site characteristics, such as the mapping of a house foundation, they are somewhat more limited in their capabilities in delineation of the site's vertical characteristics. However, as this project demonstrates, GPR does give good depth estimates and potentially allow 3D modeling and portrayal. Multiple depth slices showing structural changes with depth at Escondido and the other report examples of 3D models incorporating the vertical dimension illustrate these potentials.

The methods developed in this investigation will increasingly serve as critical steps in the evaluation of archeological properties as required by the National Historic Preservation Act. Use of these methods can increase the effectiveness and (often reduce the cost) of the evaluation efforts. Since excavation of entire sites or settlements, or even large areas of them, is impossible owing to funding limitations and ethical concerns, it may be only through integrated remote sensing that real understandings of the content, structure, and extent of archaeological sites may be achieved. It is anticipated that the methods pioneered here provide an important step in the direction of that goal.

2. OBJECTIVE

In response to CSSON-02-01, the Center for Advanced Spatial Technologies and the Department of Anthropology, University of Arkansas, proposed *New Approaches to the Use and Integration of Multi-Sensor Remote Sensing for Historic Resource Identification and Evaluation*. The focus of the proposal was the identification of specific combinations of remote sensors and data integration methods for the detection, identification, and interpretation of cultural resources in various environments and archaeological circumstances. This report presents the findings of that research.

2.1. CULTURAL RESOURCE MANAGEMENT DETECTION & EVALUATION TECHNOLOGIES: THE SERDP STATEMENT OF NEED

Kenneth L. Kvamme, University of Arkansas

Some of the nation's most significant historic and prehistoric cultural resources are contained within the 25 million acres of public land administered by the Department of Defense. Protecting these heritage resources is a fundamental part of the Department's primary mission. Heritage management issues central to that mission have focused on the economics of identifying and maintaining historic facilities, the impact of archaeological sites on construction and training programs, and the disposition and curation of artifacts. Cultural resources, defined in DoD Instruction 4715.3, include buildings, structures, sites, districts, and objects eligible for, or included in, the National Register of Historic Places regulation (36 CFR 60). Management of these resources, in compliance with existing laws and regulations such as *Native American Graves Protection and Repatriation Act* (25 USC §3001), the *Archaeological Resources Protection Act* (16 USC §470 aa-ll), *American Indian Religious Freedom Act* (42 USC §1996), and the standards in the *Curation of Federally-Owned and Administered Archaeological Collections* (36 CFR 79), necessitates the development of innovative and cost-effective methods for archaeological site identification, evaluation, and protection.

The SERDP Statement of Need (SON) entitled "Cultural Resources Management Detection and Evaluation Technologies" (CSSON-02-01) requested research into a range of advanced archaeological methods. Specifically, it called for methods that (1) "effectively detect, locate and identify historic and pre-historic archeological resources on military and DoE lands and ranges," (2) "improved models for predicting the location of resources," and (3) "improved technologies for detecting surface and/or subsurface resources." Non-invasive processes and procedures were "strongly encouraged in order to reduce the possible disturbance of human remains and associated artifacts." Finally, validation of findings through excavation was deemed "necessary to demonstrate the feasibility of the proposed technologies and associated procedures."

This Statement was timely, because mainstream archaeological methods for the identification and evaluation of most historic resources remain little changed from those employed in the early twentieth century. Surface survey and excavation, the traditional field methods for discovery of artifacts, architectural elements, and other archaeological features, continue to predominate in spite of the fact that these techniques are extremely time consuming, expensive, and unreliable. Their shortcomings manifest themselves in many ways. Frequently there is no discernable surface evidence of buried archaeological

features, making surveys ineffective. Excavation of small shovel test-pits can sometimes locate archaeological sites, but this method entails substantial additional survey costs. In National Register of Historic Places (NRHP) eligibility assessments, ever-increasing labor costs restrict the number of units that can be excavated, resulting in failures to locate features of significance. More importantly, traditional invasive methods regularly lead to the damage or destruction of the very resources they were designed to investigate and generate costs associated with collection curation. Aware of these limitations, archaeologists have adopted a conservative, preservationist approach, with the result that the Department of Defense (DoD) protects many sites of marginal scientific importance. It is apparent that the SERDP SON implicitly recommends *archaeological remote sensing* as a primary vehicle for meeting current needs.

2.2. REMOTE SENSING FOR ARCHAEOLOGICAL DETECTION AND EVALUATION

Kenneth L. Kvamme, University of Arkansas

2.2.1. Remote Sensing and Archaeology

Remote sensing, as a group of investigatory methods, is relatively new to archaeology, but it has begun to make impacts on the goals and practice of the discipline. As the planet's exploding population results in massive changes to the landscape, the need has grown for efficient and cost-effective methods to locate, map, and acquire information from sites of our cultural heritage before they are lost. At the same time, traditional archaeological excavation and pedestrian survey costs have skyrocketed, and they commonly examine only trivial areas considering the sizes of many archaeological sites, settlements, and cultural landscapes. Compared to traditional methods, all remote sensing techniques allow large regions to be rapidly investigated at relatively low cost. These methods can detect archaeological elements unseen on the surface, precisely map them, and offer interpretations based on form, distribution, context, and measurement characteristics. It is increasingly being realized that remote sensing may offer the only pragmatic means for locating, mapping, and inventorying much of the world's archaeological resources. Although this implies that remote sensing is only a descriptive tool, it also offers many methodological avenues for the *interpretation* of that record, a fact often emphasized in this study.

Hesse (2000:35) defines archaeological remote sensing as "the recognition, description, and interpretation of traces of the past, either on the surface or detectable from the surface." It can be liberally viewed as "any technique that acquires information through indirect means" (Kvamme 2005:424) which allows the ground-based methods of geophysics (e.g. magnetometry, ground-penetrating radar), aerial survey techniques (e.g. photography, thermography), and satellite imaging to be included under its rubric. This is the view taken in this SERDP project, where focus is confined to *terrestrial* remote sensing—the many methods and techniques for land-based detection (as opposed to marine remote sensing).

Recent advances in remote sensing and related technologies suggest the possibility that these techniques might form the basis for an alternative methodology for investigating buried cultural resources. In essence, archaeology needs methodologies for "seeing" into the ground to detect subsurface archaeological deposits and various remote

sensing methods have proven capable and cost effective for this purpose. A suite of space, aerial, and ground-based technologies has emerged in the past 30 years. Each has demonstrated its utility in identifying buried cultural resources (Clark 2000; Conyers 2004; Gaffney and Gater 2003; Hesse 2000; Kvamme 2001, 2005; Wilson 2000). Some technologies record variations only at the surface; others can penetrate below to a depth that varies with soil properties, vegetation cover, and types of archaeological remains (usually in the uppermost 1-2 m in archaeology). Regardless of method, useful results occur when *contrasts* or different physical properties are detected between archaeological features and the natural geological background. A buried stone foundation might be more magnetic, better reflect radar energy, more slowly emit thermal energy, or stunt overlying plant life, compared with the surrounding earth. Such contrasts are referred to as "anomalies" until their sources are identified.

2.2.2. Pattern Recognition and Large-area Surveys

Anomalies detected via remote sensing frequently illustrate sufficiently clear pattern in form or distribution for *direct* interpretation, as when the rectangle of a house foundation is unambiguously expressed (Figure 2.1). This notion forms the "pattern recognition" basis of interpretation, one of several key interpretive tools (Kvamme 2006b). It is derived from a fundamental tenet of remote sensing (Avery and Berlin 1992:52) that patterned geometric shapes in the landscape—circles, ellipses, squares, rectangles and straight lines—are generally of human origin. They occur much less frequently as products of nature. Making use of this principle demands large-area surveys. Although not a problem for space or aerial imaging methods that rapidly image square kilometers, it can be an issue with ground-based geophysics, but new and faster instruments and field methods allow surveys approaching a hectare per day (Gaffney and Gater 2003; Gaffney et al. 2000). Remote sensing of small areas might reveal anomalies with a likely cultural source, but without seeing their full forms and possible associations with other anomalies, it is difficult or impossible to interpret what they represent or ascribe significance to them. A small linear anomaly could represent part of a room, a house wall, a ditch edge, a pavement, trail, or road, for example, but a small-area glimpse may not be sufficient for secure interpretation. Surveys of large contiguous areas increase the likelihood that shape and size characteristics, plus associations with other identified anomalies, may permit recognition and identification of what an anomaly represents. In so doing, greater sense can be made of remotely detected patterns in cultural sites and landscapes (Kvamme 2006b).

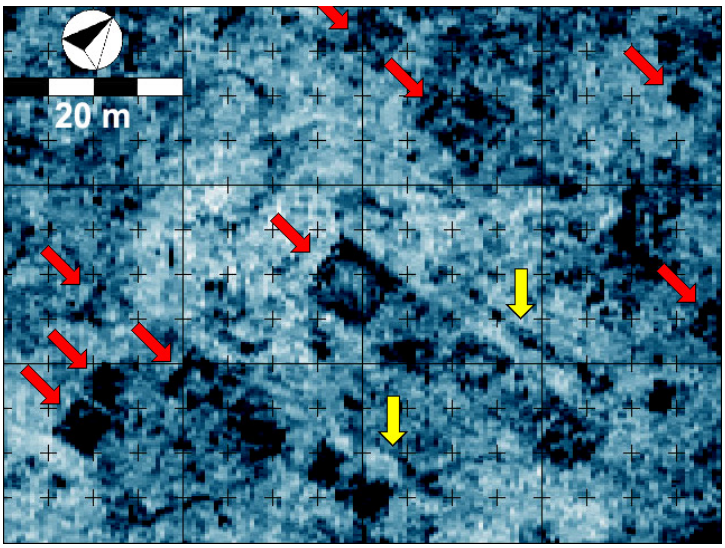


Figure 2.1. *Regular geometric shapes in anomaly forms point to cultural constructions in this ground-penetrating radar image from Pueblo Escondido (AD 1200-1400), New Mexico, one of the SERDP Project sites. Square-shaped anomalies identify Jornada Mogollon pithouses (red arrows), many arranged in rows. Certain lineations (yellow arrows) appear to be constructed features associated with the pithouses.*

2.2.3. Benefits of Remote Sensing

The benefits of remote sensing surveys to archaeology have traditionally been viewed as follows: (1) they can be employed to discover new archaeological sites and internal site features, (2) guide expensive excavation and testing programs to features of potential archaeological interest, (3) allow cost savings by making site explorations more efficient, and (4) they are non-destructive, permitting the resource to remain intact (an important consideration when exploring culturally sensitive burial, sacred, or ceremonial sites, or Traditional Cultural Properties) (Weymouth 1986; Wynn 1986). The remote detection of cultural features can reduce the amount of excavation needed, and therefore costs, to effectively evaluate National Register eligibility—because field teams can go directly to features indicated. Smaller artifact collections may also be recovered, decreasing curation charges, because the volume of excavation is reduced when archaeological features are reliably located. Moreover, the detection of subsurface features lessens the risk of inadvertent discoveries of cultural resources and attendant mitigation costs (Hargrave 2006).

The very success of remote sensing in the foregoing areas has led to the conception that these methods are suited only for discovery purposes, but recent advances and practices have allowed these technologies to evolve into new domains of inquiry. In particular, *remote sensing alone can yield primary data suitable for the study of cultural forms within archaeological sites and landscapes*. In other words, culturally patterned anomalies identified in wide-area remote sensing surveys can reveal organization and structure within settlement or larger spaces. The identification and examination of relationships between such individual site components as houses or house clusters, lanes, dumping grounds, public structures, storage facilities, gardens, plazas, fortifications, cemeteries, and the like, can be made through interpreted imagery. Likewise, inter-settlement comparisons of form, estimates of numbers of houses, average house sizes, or the examination of house shapes, orientations, and arrangements of interior components can also be attempted through remote sensing alone (Bales and Kvamme 2005; Gaffney et al. 2000; Kvamme 2003). This capability offers the possibility of direct study of settlement content and form through remote sensing. The compilation of libraries of site

plans of Neolithic enclosures and Roman villas in Britain, derived entirely from geophysical data sets, exemplifies this capability (*The English Heritage Geophysical Survey Database*, http://www.eng-h.gov.uk/SDB).

This capability has profound implications when contrasted with traditional archaeological excavations. Costs and time requirements of excavation have meant small exposures limited in extent to only a few square meters. This necessity has forced a mindset onto the practice of archaeology that commonly views human spaces in terms of tens of *square meters* as opposed to the tens of *hectares* within which most humans live and interact. In other words, human activities occur within settlements and cultural landscapes covering hectares or even square kilometers. Until remote sensing, there has not been a means to visualize subsurface distributions of cultural features over such large areas, and there is generally a tremendous amount of archaeological ignorance about the size, structure, and layout of settlements, inter-settlement relationships, and cultural uses of landscapes in most regions and cultures. Areas of the subsurface "exposed" by remote sensing are several orders of magnitude larger than can be achieved by excavation, with the possibility of commensurate increases in overall knowledge gained. These aspects alone pose the possibility of a major revolution in the archaeological knowledge base through remote sensing.

2.2.4. The Nature of Archaeological Sites

It is prudent to examine the nature of the archaeological record to understand exactly what remote sensing can and cannot accomplish. With more than a century of scientific excavations a great deal is known about the internal character of archaeological sites. This knowledge can be applied to determine the kinds of cultural features that might be remotely detected and the nature of sensor responses (see Scollar et al. 1990:4-7). It is a well-known dictum that an archaeological site is a three-dimensional matrix of materials, containing artifacts, floral and faunal remains, human-generated constructions (features), and the deposits in which they lie (see Schiffer 1976).

- *Artifacts* are material objects modified by people. Portable artifacts include small items that are easily moved, such as tools employed in day-to-day activities (e.g. spear points, pots, knives; Figure 2.2a). Non-portable artifacts include such larger items as cut posts, building timbers, bricks, or shaped stones used in architectural constructions.
- *Constructions* include the many types of buildings and other structures that people make, including places of occupancy (dwellings, public facilities), non-occupancy (storage pits, wells, burial mounds, fortification ditches), and transportation facilities (roads, trails). While many are composed of multiple non-portable artifacts such as bricks, or are made of native stone (Figure 2.2b), most are represented only by subtle changes in deposits. This is because ditches, house pits, or storage pits become filled with sediments and buried wooden structures decompose into soil.
- *Sediments and soils* are the deposits within which artifacts and constructions lie. Most result from natural processes, but many are anthropogenic or are altered by human activity (Figure 2.2c). *Additive deposits* occur where materials are accumulated by people, such as refuse dumps (*middens* containing food waste, bones, discarded portable artifacts, ash from fires) or where earth materials are

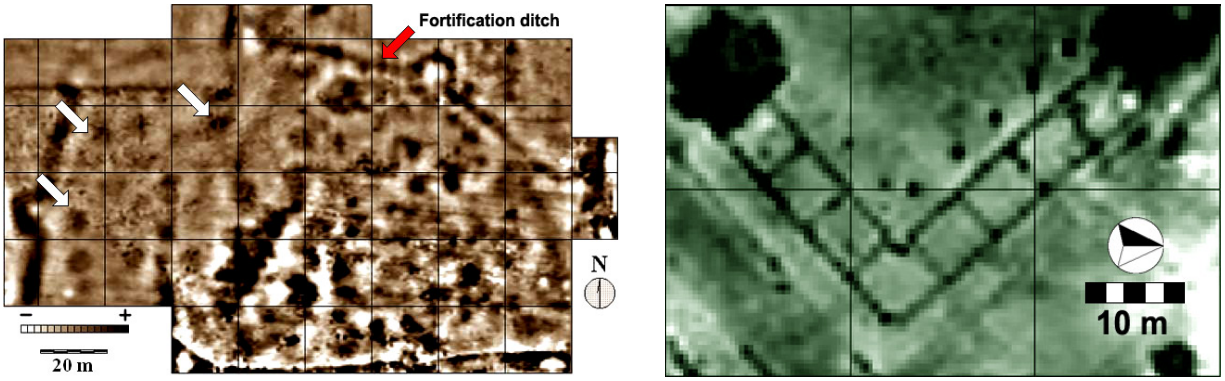
used in constructions (burial mounds, platforms, prepared house floors, raised berms). *Deposit subtraction* occurs where sediments and soils are removed, as when ditches, pits, or cellars are excavated or when incisions in the surface are caused by foot traffic. *Altered deposits* also result from human activity. Intensive firing in hearths, kilns, or burned houses can profoundly increase soil magnetism, and even the simple act of human occupancy may subtly raise it. Soil compaction commonly results over living floors or along pathways through simple foot traffic.

a **b** **c**

Figure 2.2. *Components of the archaeological record. a) Broken ceramic sherds representing portable artifacts. b) An historic pavement constructed of native stone. c) Soils and complex sediments of anthropogenic origin form stratigraphy in archaeological deposits.*

In addition to the foregoing, archaeological sites in North America are loosely divided into two principal categories (although the boundary between them remains fuzzy and the categories overlap). *Prehistoric sites* are Native American sites occupied prior to European contact, where that date varies across the continent. North of Mexico, they are generally devoid of metallic artifacts (with some important exceptions, mostly of copper and occasional meteoric iron), and stone was not generally employed for building (with notable exceptions as in the masonry Pueblos of the Southwest). Non-perishable artifacts typically consist of chipped or ground stone, ceramic, or bone. Perishable artifacts of wood, textiles, or leather are rare. Settlements are often not well structured in their layouts (Figure 2.3a). Most structural features are indicated only by changes in soils.

Historic sites are more recent and typically reflect occupations of people who came to the continent after Columbus or their descendents. It is emphasized that many historic Native American sites also exist in this late period, but they tend to share traditional characteristics of prehistoric sites with features common to Euroamerican and other historic sites. Historic sites tend to contain metallic artifacts of various types, and glass is common in addition to stone, bone, and ceramics (with china and glazed wares common). Many structural features are of stone or brick and cellars are common, although many are also indicated only by more subtle changes in soils. There tends to be a high level of geometric patterning associated with settlement layouts, for example, rectangular grids of blocks or streets, and in the form of constructed features (Figure 2.3b).



A b

Figure 2.3. *Prehistoric and historic archaeological sites revealed by remote sensing. a) Whistling Elk Village, South Dakota (ca AD 1300), showing haphazard internal distribution of houses (irregular circles shown by white arrows). b) High geometric patterning of rooms within a structure at Army City, Kansas (1917-1921), a SERDP project site. Both data sets are by electrical resistivity.*

2.2.5. Properties of Archaeological Sites: What Can Be Remotely Detected?

Large, non-portable artifacts, such as building foundation or pavement stones, generally can be remotely sensed, but the same may not be said of small portable artifacts (an exception is metallic artifacts that may be located with metal detectors). Large masses of small portable artifacts or faunal remains *might* occasionally be revealed by remote sensing when a concentration of ceramic sherds subtly raises the local magnetic field or a large amount of bone from a mass kill changes subsurface electrical, thermal, or surface vegetation patterns. Human constructions receive focus in remote sensing because their larger volumes of contrasting physical properties increase detection probabilities relative to common sensor resolutions and sampling densities (their regular geometric shapes also make them more easily recognized through the pattern recognition principle). Moreover, additive deposits are generally very different in character. Middens, filled with refuse, might contain more air voids, inclusions of solids like bone and discarded artifacts, more soil nutrients, or different levels of moisture, altering their electrical, thermal, or surface vegetation patterns (Scollar et al. 1990; Weymouth 1986). Mound creation often employs topsoil (which generally exhibits high magnetic susceptibility), raising the local magnetic field, while areas where topsoil was removed for cellars or ditches tend to locally deflate it. Heavily fired features, including kilns or burned houses, are particularly detectable because they profoundly increase soil magnetism (Clark 2000). Large constructions near the surface can impact patterns of surface vegetation, as when a buried wall of stone retards growth or moist sediments filling a former ditch advance it, forming the “crop mark” phenomenon. In either case the spectral properties of the plants may be altered (Scollar et al. 1990; Wilson 2000). Variations in deposit materials, their compaction, moisture retention and other factors affect absorption and emissive rates of solar radiation; when this energy is re-radiated thermal variations can also occur (Dabas and Tabbagh 2000).

2.2.6. Sensor Mixes and Detection Probabilities

The foregoing suggests the likelihood of detecting a buried archaeological feature depends on (1) matching the contrasting physical properties of the feature with sensors capable of detecting those properties, (2) the amount of physical contrast between the feature and surrounding deposits with respect to sensor sensitivity, (3) the size of the feature relative to the spatial resolution of the measurements, (4) the depth of the feature below the surface with respect to signal attenuation and the level of confounding noise factors that might overly it, (5) the degree of regular pattern the feature exhibits (the pattern recognition property), and (6) the use of multiple sensors that allow detection of different physical consequences of buried materials or deposits.

The spatial resolution of measurements determines the size of archaeological features that can be resolved. It should be obvious why larger features like houses or ditches are relatively easy to locate (many more measurements might define them), and why resolving small portable artifacts and is so difficult (sampling densities rarely are made at the centimeter level, the size of most artifacts). As noted previously, ground-based and aerial methods generally yield high spatial detail, well into the sub-meter range, and recent satellite systems now deliver a similar capability with meter-level spatial resolutions (Fowler 1996; Kennedy 1998). Detection probabilities are also affected by instrument sensitivity. A magnetometer with a precision of 0.001 nT can detect smaller, deeper, and more subtle anomalies than one that measures only to the 0.1 nT level, for example (Becker 1995).

Environmental conditions at the time of data acquisition also influence results. Daylight is necessary for many forms of air or space imaging, but may be undesirable for thermal remote sensing. Low sunlight is necessary to achieve terrain shadowing (necessary to detect relief changes) in aerial photography. Seasonal effects must also be considered because vegetation at a particular stage of development is necessary for "crop marking," visible from the air or space, and resistivity-conductivity methods cannot be explored in frozen ground. Too much or too little soil moisture can negatively affect soil resistivity, conductivity, or thermal surveys, while too much moisture may impede transmission of radar energy. Deeply buried archaeological features are more difficult to detect than shallowly buried ones because the soil above acts like a filter that degrades the signal from lower levels. Buried pipes, electrical lines, radio or cell phone transmissions can negatively impact many ground-based instruments. In general, remote sensing of any kind is more successful in open fields with uniform ground cover, while heavily vegetated, wooded, or urban landscapes can pose great difficulties or make many methods of remote detection impossible (Bewley 2000; Clark 2000; Dabas and Tabbagh 2000; Scollar et al. 1990).

2.2.7. Prognosis

With this background, the following observations can be made concerning remote sensing in its role as an archaeological discovery and evaluation tool, and the prognosis is good.

1. *Many archaeological sites are very large.* Remote sensing methods that offer rapid coverage are necessary for cost-effective reconnaissance of large areas.
2. *Archaeological sites, settlements, and landscapes exhibit spatial organization***.** Pattern recognition of cultural forms within remotely sensed data sets is important

and critical to interpretation. Such dimensions as shape, size, arrangement, orientation, and context must be considered, and the large-area coverage of remote sensing meets this need.

3. *Cultural features within archaeological sites possess high spatial frequency.* Many cultural features like storage pits, privies, post molds, or hearths are on the order of a meter in size or smaller. Larger features like house pits or middens (trash areas) are larger, on the order of 4-15 m in diameter. Other linear features like trails, roads, or ditches (e.g., fortification or encircling ditches) may be many meters long in one direction, but it is their short-axis dimension, from sub-meter to perhaps 1-5 m, that is relevant to remote sensing. We can infer that successful remote sensing demands very high spatial resolutions, preferably at the sub-meter level to successfully detect the small features that principally comprise an archaeological site.
4. *Soils, sediments, and other materials within archaeological deposits possess different physical properties.* Anthropogenic enrichment of topsoil, past removal or mounding of topsoil, burning, the importation of sediments (e.g., clays for floors) or rock (for floors/foundations), and the addition of ferrous artifacts beginning in proto-historic times alter magnetic susceptibilities of deposits and the cumulative magnetic field. Variations in material, porosity, particle size, salinity, moisture, and other characteristics of deposits impact their conductivity and dielectric properties. Geophysical instrumentation exists for the measurement of these various dimensions, in some cases at specific target "depths" below surface.
5. *Soil and moisture changes within near-surface archaeological deposits influence surface vegetation patterns.* The crop-mark phenomenon, well studied in European archaeology (Wilson 2000), illustrates the fact that plants growing immediately over buried stone walls or pavements, for example, will tend to exhibit stunted growth or yellowing. Other plants growing over wet, organically rich sediments that may have in-filled a ditch or house depression will tend to exhibit enhanced growth. That such differential growth patterns are most readily detected in the visible or near-infrared portions of the electromagnetic spectrum has profound implications for multispectral aerial and satellite remote sensing. Related soil and moisture changes may also be detectable in the mid-infrared range.
6. *Soil type, soil density, and moisture changes within near-surface archaeological deposits together with surface vegetation patterns yield thermal variations recordable at the surface.* Patterns in archaeological deposits can be expressed through minute temperature changes (e.g., 0.1^{o} C) at the surface. Under proper conditions, airborne thermal sensors can record these variations.

2.3. PROPOSED RESEARCH: INTEGRATED ARCHAEOLOGICAL REMOTE SENSING FROM LAND, AIR, AND SPACE

Kenneth L. Kvamme, University of Arkansas

In response to CSSON-02-01, the Center for Advanced Spatial Technologies and the Department of Anthropology, University of Arkansas, proposed *New Approaches to*

the Use and Integration of Multi-Sensor Remote Sensing for Historic Resource Identification and Evaluation. The focus of the proposal was to identify specific combinations of remote sensors and data integration methods best suited for detecting archaeological resources in a range of environmental circumstances. The goals of this research include:

1. Application of a suite of established and newly fielded high-resolution space and aerial instrumentation to the detection of cultural resources.
2. Examination of a suite of existing and newly developed ground based instruments applied to the detection and evaluation of subsurface deposits.
3. Application of existing and the development of new computational approaches for the integration of an extensive suite of satellite, aerial, and ground-based remote sensing data sets.
4. Assessment of the results of these approaches through an archaeological field validation program in a range of environments and archaeological site types.
5. Evaluation of the effectiveness of various sensor combinations and fusion methods against site conditions, and development of guidance documents for their use in various settings.

Although a single sensor can provide great insights about subsurface archaeological conditions (as in Section 2.2), it is shown that expanded benefits are realized through the integration of data from *multiple* sensors. Numerous data integration and fusion methods developed by this project effectively combine the information content from multiple sensors, yielding a dramatic increase in the ability to detect buried archaeological features, recognize cultural patterns in the data, and develop archaeological interpretations. The combination of multiple sensor systems, many of which are relatively new technologies, together with new computational approaches to multi-sensor data integration, fusion, and analysis, set the stage for the development of significant new capabilities for identifying and evaluating archaeological resources. As a result, elements of this project may represent a significant turning point in the practice of archaeological field projects.

2.3.1. Data Types Investigated

This project examines the utility of an unprecedented range of archaeological remote sensing methods for identifying and evaluating buried cultural resources through their integration. Ground-based geophysical, aerial, and satellite technologies are employed for this task. The methods investigated include (1) magnetometry, (2) magnetic susceptibility, (3) electrical resistivity (4) electromagnetic conductivity, (5) ground penetrating radar, (6) aerial thermal infrared, and (7) high resolution multispectral imagery acquired from satellites. Table 2.1 presents a summary of these methods, their data formats, and the phenomena to which they respond.

Table 2.1. *Basic Data for Instruments and Sensors Employed in Project.*

Instrument/ Sensor	Domain	Spatial Resolution	Sensing depth	Units & Precision	Data Format	Areal Coverage/ day	Type of Physical Phenomena Sensed
Magnetic Gradiometer	On ground	.125-.25 m	1-1.5 m	.1 nT	Floating point	.5-.8 ha	Magnetic field gradient
Magnetic Susceptibility	On ground	0.25-1.0 m	0.5 m	.01 ppt	0.25-1.0 m	.5-.8 ha	Induced soil magnetism
Electrical Resistivity	On ground	0.5-1.0 m	2 m	.01-1 ohm	Floating point	.4-.6 ha	Soil resistivity
EM Conductivity	On ground	0.25-1.0 m	1.5 m	.1 mS/m	Floating point	.5-.8 ha	Soil conductivity
GPR	On ground	0.02 -.5 m	2-3 m	1 nS	16 bit	.1-.3 ha	Soil dielectric changes
Thermal IR	Aerial	Sub-meter	0.5-1 m	0.1 C	8 bit	1-4 ha	Emitted IR
ATLAS	Aerial	5 m	0 m	15 bands 0.42-12.4 μm	8 bit / band	Sq. km	Reflected light (visible-mid infrared) & emitted thermal
QuickBird Panchromatic	Space	0.61 m	0 m	1 band 0.45-0.90 μm	11 bit	Sq. km	Reflected visible light
QuickBird Multispectral	Space	2.4 m	0 m	4 bands 0.45-0.89 μm	11 bit / band	Sq. km	Reflected light (visible-near infrared)

2.3.2. Benefits of Multiple Detection Methods and Data Integration

As indicated in Section 2.2, remote sensing includes a broad range of sensors that provide measurements of visible, near-visible, thermal, magnetic, electromagnetic, and electrical properties of the surface *and subsurface* of archaeological sites. Most respond to only a narrow range of physical properties, so use of multiple detection methods generally offers greater insights (Clay 2001; David 2001).

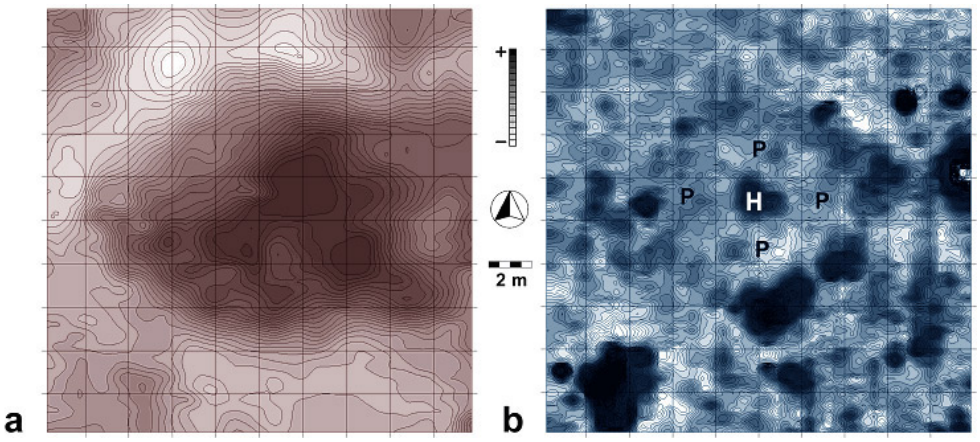


Figure 2.4. *Different geophysical methods yield complementary information. At Whistling Elk Village, South Dakota, (a) electrical resistivity data clearly define a square house floor (98 cm below surface) with a southeast facing linear entryway, while (b) magnetic gradiometry data reveal its central hearth (H), much of its perimeter showing that it was burned, and interior features including principal roof support posts (P)(all confirmed by excavation).*

Buried archaeological features not revealed by one may be made visible by another or provide complementary information. A hearth and burned elements of a buried house might be revealed by magnetometry, while electrical resistivity might delineate the outline of its compacted floor area (Figure 2.4; Toom and Kvamme 2002). Considering both data sources leads to better understanding of the form and organization of the structure.

Archaeologists commonly display multidimensional remote sensing results side-by-side, as in Figure 2.4. In an interpretation stage, outlines of selected anomalies in one data set are drawn and overlaid onto the other. Although an effective approach, it is time consuming and subjective because one is forced to make decisions about features that warrant consideration. Moreover, this approach is impractical for application to more than a few data sets because of the complexity of multiple overlays. The intent of this project is to go a step farther by integrating or "fusing" information from all sensors simultaneously into *single composite images and data sets*. In so doing, amorphous or otherwise vague anomalies that occur at a particular locus in two or three data sets might become better resolved and more interpretable. Several other benefits are also clear. A single survey, for example, might reveal only part of a buried building. Integrated information from several surveys may illustrate the entire structure as well as interior components. Moreover, integrated data may *simultaneously* show relationships between conductive, resistant, magnetic, thermal, and metallic anomalies, potentially improving knowledge of features within a site, inter-sensor relationships, and enhancing overall interpretations. In general, the use of multiple sensors and data integration permits (1) different anomalies to be revealed, yielding a more holistic view of the subsurface, (2) repetitive anomalous indications can corroborate results yielded by other sensors, and (3) more clues facilitate specific subsurface identifications. All have profound implications. By revealing more anomalies the use of several sensors augments understanding of spatial relationships between subsurface elements. The co-occurrence of the *same* anomalies indicated by different methods lends confidence, but also allows subtle discoveries barely visible by one technique to be indicated more robustly and securely. A single anomaly revealed by several sensors also offers more clues about its physical properties, giving more to work with when forming interpretations (Kvamme 2006a).

All these benefits of integrated data mean improvements in practical archaeological applications of remote sensing (Section 2.2.3). For example, (1) more archaeological features within sites may be revealed, and revealed more securely, leading to (2) greater accuracy in guiding excavations to archaeological elements of interest, and therefore (3) greater project efficiency and cost savings. Clearer imagery that *looks like* the buried archaeology facilitates interpretation by the non-specialist, from common archaeologists wishing to better understand their sites to land managers, State Historic Preservation Office (SHPO), and Native Americans involved in the regulatory process. Anomalies may potentially be prioritized by how robustly they are indicated across multiple dimensions. A greater understanding will also be possible of those anomalies indicated multi-dimensionally, allowing more informed guidance of field investigations and minimization of false positives (anomalies *not* associated with cultural deposits). Part of this reduction rests on accurately differentiating between those anomalies related to cultural features from those associated with natural phenomena (e.g. tree roots, rodent

burrows, etc.). Fused, multi-sensor data enhances the potential for this differentiation, based on visual or computer-assisted assessments of size, shape, and distribution.

There is also much more to gain in more theoretical domains through data integration. For example, the totality of subsurface evidence from all sensors may be simultaneously displayed, leading to improved pattern recognition, understanding of spatial relationships between individual site elements, and awareness of settlement layouts, organization, and structure. All are important for anthropological and archaeological understandings of past cultures and human spatial behavior. Individual archaeological features may be better resolved, improving their identifications, but also leading to other information, such as accurate area and volume estimates—of relevance to estimating the number of residents in a house to the capacity of subterranean storage pits, for example. Simultaneous portrayal of multiple remotely sensed dimensions can also improve understanding of sensor responses to specific subsurface conditions and further enhance knowledge in geophysical theory.

2.3.3. Project Objectives

A principal goal of this SERDP research project is a determination of remote sensing methods and techniques that work well individually, and that complement each other collectively when integrated, for the identification and evaluation of buried archaeological remains. Numerous methods for data integration are explored. With the exception of certain multi-band visualization techniques and overlays of vectors representing interpreted anomalies, most of the methods investigated have not been applied previously in archaeology. Several advanced computer graphic methods are explored, discrete methods that range from *Boolean* overlays to sums of categorized portrayals of sensor outputs and cluster analyses, continuous methods that include sums, products, ratios, principal components, regressions, and probability surfaces, intelligent knowledge-based systems like *C5.0* and *Cubist*, and Expert Systems approaches. This research demonstrates that certain integrating methods yield more information about the subsurface than others, but what may be realized in each approach may depend on overarching purpose. Some fusion techniques yield visually pleasing results that appear to well-combine available information, for example, while others may seem less revealing but offer greater interpretive or predictive potential. In this process, the nature of similarities, differences, redundancies, and performance characteristics of results are examined. An important aspect of this research is an assessment of the added value of the fused product compared to traditional, individual-sensor based analysis. The use of commercial off-the-shelf software solutions makes it possible to quickly transfer the results and technology to a broad user base.

The process of archaeological remote sensing as carried out in this project is a multi-step undertaking. The first stage, designed to meet the primary data integration goal, includes (1) remote sensing data collection (Section 4.2), (2) processing to clarify anomalies in individual sensors established within GIS databases (Section 4.4), (3) data fusion to integrate information from all sensors (Sections 5.2-5.11), (4) the definition of potentially "significant" cultural anomalies (Section 5.12), and (5) classifying the anomalies into likely types of cultural features (Section 5.12). The last might include indications of material composition and depth data. This is normally the end result of

most remote sensing projects, and where archaeological fieldwork takes over. This project endeavored to go several steps farther with a *second* principal goal.

The second goal of this research project was very different and designed to meet an important criterion in the SERDP Statement of Need. Specifically, "ground-truth" testing was called for to demonstrate the feasibility of the proposed data integration technologies and associated procedures. Three additional project tasks were therefore designed. They include (6) a sampling design (Section 5.14) that allows (7) archaeological excavation of representative anomalies of each defined type to provide validation of remote sensing predictions about the subsurface (Section 5.16). This validation phase often turned into a learning process, however, because the soils, geology and the archaeology in each site are unique, idiosyncratic, and confound predictability. In other words, remote sensing predictions cannot be perfect and a look into the ground through excavation offers additional insights that allow modifications to original predictions. Consequently, a final stage was defined that includes (8) modification of original remote sensing predictions, based on excavation findings (Section 5.17). Preparation of this report (9) was, of course, the ultimate task.

Table 2.2. *Principal SERDP Project tasks accomplished by season and year. Numbered items refer to components discussed in text.*

<table>
<tr><th colspan="2">2002</th><th colspan="3">2003</th><th colspan="3">2004</th><th colspan="3">2005</th></tr>
<tr><th>Summer</th><th>Fall-Winter</th><th>Winter-Spring</th><th>Summer</th><th>Fall-Winter</th><th>Winter-Spring</th><th>Summer</th><th>Fall-Winter</th><th>Winter-Spring</th><th>Summer</th><th>Fall 05-Winter 06</th></tr>
<tr><td>1) Project initiation</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td colspan="2">Equipment purchases</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td colspan="5">Aerial &satellite data acquisition</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td colspan="3">Ground-based data acquisition</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td colspan="4">2) Preprocessing of data sets; establish GIS databases</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td><td colspan="4">3, 4) Data analysis, fusion & focus on significant anomalies</td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td colspan="2">5,6) Classify anomalies; design field sampling</td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td colspan="2">7) Archaeological fieldwork & validation</td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td colspan="2">8) Project analysis, evaluation, reporting</td><td></td></tr>
<tr><td></td><td>Annual report</td><td></td><td></td><td>Annual report</td><td></td><td></td><td>Annual report</td><td></td><td></td><td>9) Final report</td></tr>
<tr><td colspan="11">Quarterly progress reports submitted at appropriate times</td></tr>
<tr><td colspan="2">PIs attend annual SERDP symposium</td><td colspan="3">PIs attend annual SERDP symposium</td><td colspan="3">PIs attend annual SERDP symposium</td><td colspan="3">PIs attend annual SERDP symposium</td></tr>
</table>

These operations were undertaken at each of four prehistoric and historic archaeological sites distributed across time and space in a wide diversity of environmental settings from South Carolina to New Mexico (Section 3.3). This allowed a

means to ascertain the value of the methods investigated at very different archaeological sites with very different remains in very different environments. In so doing a better understanding of which methods consistently worked and offered useful results could be achieved, but this knowledge was also augmented by the considerable experience of SERDP Project team members.

This multi-year project was divided into seasonal and yearly tasks, with preliminary results, conference papers, and annual reports a constant part of the cycle. Owing to budgetary delays some project tasks were deferred. The project tasks and events accomplished in this project, or "milestones," are temporally organized in Table 2.2.

2.3.4. Deliverable Products

The proposal *New Approaches to the Use and Integration of Multi-Sensor Remote Sensing for Historic Resource Identification and Evaluation* named 20 project deliverables. Nearly all are presented in the following pages, and some occur in previous reports, or in SERDP annual symposia and meetings. Section numbers of this report indicate where relevant materials may be found.

1. Description of ground-based field methods and instrumentation (Sections 4.2-4.3).
2. Description of satellite methods and instrumentation (Section 4.7).
3. Description of data processing methods (Section 4.4).
4. Graphical and data description of remote sensing findings from each study site (Section 4.6).
5. Interpretation of remote sensing findings from each study site (Section 4.6).
6. Description of data integration/fusion methods (Section 5.2-5.11).
7. Data fusion imagery from each study site (Section 5.2-5.11).
8. Operational data fusion software system. (ESTCP Project SI-0611)
9. Interpretation of integrated remote sensing findings from each study site (Section 5.12).
10. Description of archaeological field validation research design (Section 5.14).
11. Description of archaeological field validation methods (Section 5.15).
12. Description of archaeological field validation results and multidimensional remote sensing accuracy assessments (Section 5.16).
13. Discussion of kinds of archaeological features found and missed, false positives and negatives (Section 5.16).
14. Discussion of benefits of multidimensional remote sensing data integration (Section 9).
15. Discussion of data redundancies and cost-benefits within each ground, air, or space domain and between domains (Section 9).
16. Recommendations of the best subset of instrumentation necessary for future work of this kind (Section 9).
17. Discussion of environmental characteristics and their effect on detection (Section 9).
18. Annual technical reports. (Submitted to SERDP, available online at http://www.cast.uark.edu/cast/serdp/index.html)

19. Final technical report. (Submitted to SERDP, available online at http://www.cast.uark.edu/cast/serdp/index.html)
20. Presentations or posters at annual SERDP symposia and meetings. (Appendix B Technical Publications)

3. BACKGROUND

This section identifies Department of Defense regulatory requirements pertaining to cultural resources. It defines the problems and technologies addressed by this project in terms of those requirements. Past research in the problem area is reviewed and historical background material is presented describing the archaeological sites investigated by this project.

3.1. DoD REGULATORY REQUIREMENTS AND PROJECT FOCUS

Michael L. Hargrave, Construction Engineers Research Laboratories (CERL) and Kenneth L. Kvamme, University of Arkansas

3.1.1. DoD Regulatory Requirements

In this report the term Cultural Resource Management (CRM) is used in reference to the profession that has, since the mid 1960s, evolved in response to the need to comply with federal historic preservation law. All of the states, most land managing agencies (including DoD and Army), and some counties and cities have developed additional regulations relevant to CRM. Cultural resources, defined in DoD Instruction 4715.3, include buildings, structures, sites, districts, and objects eligible for, or included in, the National Register of Historic Places (36 CFR 60). For purposes of this report, cultural resources are generally synonymous with prehistoric and historic archaeological sites. In reality, cultural resources also include historic buildings and landscapes, traditional cultural properties, sacred sites, and objects.

Passage of the National Historic Preservation Act (NHPA) in 1966 was the single-most important event in the development of the CRM profession. Subsequent laws and executive orders (e.g., the Archaeological and Historic Preservation Act of 1974 as amended, the Archaeological Resources Protection Act of 1979 as amended, the Native American Graves Protection and Repatriation Act of 1990 as amended, Executive Orders 11593 and 13007) have expanded the responsibilities of federal and state agencies for managing cultural resources. One of the NHPA's most important provisions was the establishment of the National Register of Historic Places (NHRP), the nation's official list of cultural resources that are worthy of preservation. The NHPA also established the requirement for federal agencies take into account the effect of their undertakings on historic properties. Agencies must make responsible efforts to identify historic properties (most archaeological sites are essentially 'invisible' to the untrained eye), to determine if they will be impacted by proposed undertakings, to evaluate their eligibility for nomination to the National Register of Historic Places, and to seek ways to avoid, minimize, or mitigate adverse impacts to NRHP eligible properties. The process for complying with Section 106 of the NHPA is stipulated by 36 CRF 800 (Protection of Historic Properties).

The DoD administers 25 million acres of public land containing some of the nation's most significant historic and prehistoric cultural resources. Protecting these heritage resources is therefore a fundamental part of the Department's primary mission. The Army has developed regulations that specify how compliance with NHPA and other key historic preservation laws is to be accomplished (Army Regulation 200-4). Compliance with NHPA is a particularly serious challenge for the Army. Its 120 major

installations are distributed across the nation, particularly in the Southeast, Southwest, Midwest, and West. The 12 million acres managed by the Army contain some 90,000 known archaeological sites, 64,000 of which are located in maneuver areas. To accomplish its primary mission of national defense, the Army must maintain a program of realistic training. The need to train mechanized units as well as infantry, and to maintain the infrastructure needed to do so, requires the intensive, sustained use of large, contiguous land areas. Unfortunately, archaeological sites tend to be widely distributed across the landscape and represent serious impediments to effective training. Federal law requires sites that *may* be eligible for the NRHP to be treated in the same way as those that have been determined to be eligible. For practical purposes, this means that training activities must avoid many of the known sites, as well as areas where reconnaissance for archaeological sites has not been completed.

Determining an archaeological site's eligibility for the NRHP is expensive, with costs per site in many states ranging from $10,000 to $20,000. NRHP eligible sites whose presence poses unacceptable limitations on training can be "mitigated." Mitigation programs are intended to offset the adverse impact of site destruction by making scientific and cultural information about the site available to the scientific community and, to some extent, to the general public. In most cases, mitigation entails the careful excavation of a large portion of the site, detailed analyses and long-term curation of the artifacts, and the preparation of a scientific report. Site mitigations are generally expensive, with costs easily exceeding $100,000 for moderately large or complex sites.

Archaeological sites generally consist of three types of deposits: 1) artifacts (tools, manufacturing debris); 2) sediments enriched by organic residues associated with human occupation (ash, bone, food remains, etc.), and 3) discrete constructed features (architectural remains, pits, hearths, graves, etc.) (see Section 2.2.4). Such features are generally important to an assessment of a site's NRHP eligibility status because they sometimes contain artifacts, food residues, and datable materials that represent activities conducted during relatively brief time intervals. Archaeological sites that lack discrete subsurface features are sometimes found to be eligible for the NRHP, and sites that contain features are often found to be non-eligible. Information about the presence, condition, and contents of subsurface features is, however, almost always necessary in order to assess NRHP eligibility. Archaeologists in the US generally use several different field methods to collect the data needed to characterize sites, evaluate their eligibility for the NRHP, and to recover data needed to mitigate the effects of site destruction.

DoD efforts have therefore focused on the economics of identifying and maintaining historic facilities, the impact of archaeological sites on construction and training programs, and the disposition and curation of artifacts. Management of these resources, in compliance with existing laws and regulations necessitates the development of innovative and cost-effective methods for archaeological site identification, evaluation, and protection.

3.1.2. Project Problem Focus and Regulatory Requirements

As outlined in Section 5, this SERDP Project, *New Approaches to the Use and Integration of Multi-Sensor Remote Sensing for Historic Resource Identification and Evaluation*, focuses on ways to combine remote sensing data from multiple land, air, or satellite sensors to detect, describe, and interpret subsurface cultural resources in

archaeological sites. The effective management of cultural resources requires knowledge of exactly *what* resources lies beneath the ground and *where* they are located. Moreover, in order to make assessments for National Register eligibility, some idea of the significance of those resources must be established. The combination of multiple sensor systems with new computational approaches to multi-sensor data integration offers significant new capabilities for identifying and evaluating archaeological resources. This SERDP Project offers new in-roads in these domains.

3.1.2.1. Deficiencies of current archaeological discovery methods

The methods generally employed for identifying and characterizing archaeological sites unfortunately all possess serious shortcomings that make subsurface detection and discovery unreliable (Kvamme 2003).

Surface survey methods by pedestrian inspection of a site's surface can only recover uneven samples of artifacts from the surface, and only in freshly plowed fields or in arid landscapes that lack surface vegetation cover. Artifact collections are typically heavily biased by factors that affect their visibility, including object size and color. Architectural features and other constructions are rarely discovered and located unless elements protrude to the surface or significant alterations of the topographic surface occur. These features can rarely be precisely located based solely on surface artifact distributions.

Shovel Test Pits (STPs) are small excavated holes about 30 x 30 cm in area and perhaps 50-70 cm deep that are typically excavated at 15-30 m intervals in an attempt to locate evidence of buried archaeological sites, characterize artifact densities and distributions, and search for subsurface culturally constructed features. They are the primary method used by American archaeologists for subsurface prospecting, and staggering numbers of them have been excavated across the country at great cost, despite low discovery probabilities. Shott (1985:457) shows that STPs "performs poorly as a site discovery technique" and Krakker et al. (1983:475) conclude that STP sampling can discover only "a very small percentage" of sites. In the Ozark-St. Francis National Forest in Arkansas, thanks to careful record keeping (Jurney 2001), it is estimated that approximately 40,000 sterile holes have been excavated, for example, with fewer than 400 yielding positive evidence of cultural remains (a "hit rate" of only one percent). Moreover, in about 85 percent of the sites discovered so little information was gained that National Register determinations could not be made (Jurney 2001:9). Within known archaeological sites widely-separated STPs pose similar problems by failing to encounter small, widely spaced archaeological features that are essential to an assessment of site significance and NRHP eligibility. Their inability to reliably locate features of small size, such as graves, can lead to costly mitigations.

Controlled hand excavations (e.g., 1 x 1 m units excavated in 10 cm levels using shovels and trowels) are expensive and time consuming. This technique offers the best opportunity to document the details of stratigraphy within an archaeological site, recover artifacts, floral, and faunal remains, and document characteristics of architecture and other constructed features. The slowness of the process and exactness of modern recording techniques permit the highest level of archaeological knowledge to be gained—but only within the confines of very small spaces. Only small exposures of a few tens of square meters typically are made, and large archaeological sites may encompass

thousands of square meters. This result means that site significance must be evaluated based on the investigation of tiny samples (often less than 1-2 percent) of each site.

Mechanized excavation is sometimes used to remove the uppermost plow-disturbed stratum to expose subsurface archaeological features that lie immediately below. This approach can be effective at open, shallow sites, but it is highly invasive, and creates a risk of encountering near-surface human remains. Mechanized excavation is thus not appropriate at sites likely to include graves, complex deposits, or sites that are located in forested areas.

Invasive excavations characterize all traditional archaeological methods of detection and discovery, except the first, meaning that part of the archaeological resource base is destroyed in the recovery of information. Ground-disturbing excavations in culturally sensitive burial, sacred, or ceremonial sites, or in Traditional Cultural Properties, may be unacceptable, leaving often-unrevealing surface inspection methods as the only traditional approach for investigation.

3.1.2.2. This project's benefits and regulatory requirements

The methods proposed in this project, based on the integration of information from multiple remote sensing devices, confront many of these difficulties. (1) Remote sensing methods are non-invasive. The resources remain intact, causing no adverse impacts to human remains, sacred sites, and other sensitive deposits. (2) Large areas can be surveyed at very high data densities. This potentially allows a wide range of subsurface archaeological features and deposits to be revealed and mapped (but see Section 5.12.2). In other words, it is possible to image aspects of the subsurface of entire archaeological sites and settlements. (3) Because the loci of subsurface archaeological features can be revealed, excavations needed to document stratigraphy and depositional integrity, and to recover artifact samples and datable materials can be placed *exactly* over those features (Johnson and Haley 2006). The net effect is a reduction in the amount of excavation that may be required in a project (since their placement can be guided and they need not be haphazardly placed), greater cost efficiency, and smaller collections of archaeological artifacts. (4) The *integration* of information from multiple sensors, the principal emphasis of this project, means *greater information content*. The implications are that: (a) more subsurface features are potentially revealed than in traditional remote sensing projects that employ only a single sensor, so more of the subsurface is "exposed"; (b) complementary information yielded by multiple sensors corroborates findings, making them more secure; (c) information representing several physical dimensions of the subsurface gives more clues about the identity, make-up, and content of buried archaeological entities, *allowing more accurate interpretations*.

Lockhart and Green (2006) argue that remote sensing represents a kind of "preservation in place" because the partial recovery of archaeological data is obtained while the resource remains intact. Kvamme (2003) goes a step further by asserting that remotely sensed information can serve as *primary* data suitable for archaeological analyses of site form, content, distribution, and structure in those cases where imagery is particularly clear. Remote sensing offers other benefits to the regulatory process. In consultations with non-specialists, such as members of the public, SHPO, and Native Americans, for example, clear and interpretable imagery of the subsurface can facilitate communication about the nature, number, size, and distribution of archaeological

features. This kind of information can be more representative of a site's content and more clearly convey the nature of the subsurface than complex discussions of stratigraphy and site plans gained by traditional means that are typically from investigations much smaller in area (Lockhart and Green 2006).

3.2. HISTORY OF DATA INTEGRATION IN ARCHAEOLOGICAL REMOTE SENSING

Kenneth L. Kvamme, University of Arkansas

In satellite remote sensing the integration of multi-sensor, multi-band data has been commonplace since its inception. Even the earliest satellites, such as LANDSAT 1-4, acquired multi-spectral information from bands ranging through visible to near infrared. Their integration through ratioing, RGB color compositing, and principal components analysis was routinely undertaken, because it was assumed, a priori, that multidimensional information was much richer in content (Moik 1980; Schowengerdt 1983). Resulting data composites proved richer and more informative than any single input band and these products became de facto standards. By the mid-1980s, remote sensing scientists were going a step further by effectively integrating satellite information with ancillary data, particularly digital terrain models and such derived products as slope and aspect, in efforts that substantively improved accuracies of image classification efforts (summarized in Hinton 1996).

Archaeologists have long focused on diverse data sets in their investigations of archaeological sites. Maps of surface artifact distributions (Sullivan, ed., 1998), excavation finds (Renfrew and Bahn 2003), historic and recent aerial photography (Wilson 2000), geophysical mappings (Gaffney and Gater 2003), and even aerial multispectral and satellite imagery (Fowler) have been examined along with standard topographic maps and thematic soils, geology, and other environmental maps, to better understand site content, structure, and context within landscapes. Typically, even with the availability of such diverse information, results were only considered side-by-side as distinct mappings, although cross-correlating information was certainly undertaken. The advent of GIS technology in the late 1980s certainly promoted the integration of archaeological information from diverse inputs because it became easy to overlay data and generate composite maps (Kvamme 1989), if only because GIS demands a common spatial coordinate base that forces such data linkages (van Leusen 2001). Discrete entities observed in one map (in digital or paper form) might be combined with discrete mappings in another to produce some sort of data composite. This practice allowed van Leusen (2001:575) to observe that when several archaeological data sets are integrated the "whole…is larger than its constituent parts." Similarly, David (2001:525) asserted that such contexts offer a maximum of information about a site non-destructively, and meet a critical management need.

In archaeological applications of geophysical surveys, Weymouth (1986:371) observed that multiple geophysical methods *must* offer improved insights because each yields information about a relatively independent aspect of the subsurface. In a similar vein, Clay (2001) effectively demonstrates why "two techniques are better than one." An early effort that actually practiced data integration was the *Wroxeter Hinterlands Project* (Gaffney and van Leusen 1996). They combined a full-scale magnetic gradiometry

survey of this important Romano-British town (about 0.8 km^2) with more limited electrical resistivity data and a host of vertical and oblique aerial photographs that showed outlines of structures through vegetation marks (oblique photos were rubber-sheeted to the common coordinate base). This allowed them to digitize interpretations in individual data sets and overlay them to yield composite interpretations for the entire site. Brizzolari et al. (1992:54), Doneus and Neubauer (1998), Buteux et al. (2000:77), and many others, have presented overlaid interpretive vectors representing cultural anomalies revealed by multiple geophysical and aerial surveys.

It is only in recent years, however, that archaeological data integrations more complex than simple overlays of interpreted vectors have been generated. Neubauer and Eder-Hinterleitner (1997) and Piro et al. (2000) advanced a suite of simple mathematical operations that effectively combine geophysical data. Using sums, products, or ratios between different inputs, their results offered *mathematical* integrations of archaeological data for the first time. Kvamme (2001) presented a computer graphics solution where a color composite of three geophysical data sets fused findings within a single image. Johnson and Haley (2004) pioneered supervised classification methods to better define subtle anomalies expressed in multidimensional geophysical data. In each case, patterns were enhanced because information from multiple inputs was simultaneously combined in one result.

It is these latter groups of approaches that this project attempts to utilize and further develop. Several goals in this work are apparent. One lies in developing superior ways to visualize the totality of archaeological evidence by portraying multiple dimensions simultaneously. There is also a desire to improve theoretical understandings of relationships between various data sources and the hope to more achieve more accurate interpretations of the subsurface. With little effort thus far expended in synthesizing available approaches to data integration, this project hopes to formalize it as a legitimate field of archaeological study.

3.3. SELECTING PROJECT SITES

Michael L. Hargrave, Construction Engineers Research Laboratories (CERL), with a contribution by Eileen G. Ernenwein, University of Arkansas

The SERDP project provided a nearly unique opportunity to conduct large scale, multi-sensor surveys and subsequent ground-truthing investigations at four archaeological sites. The only restriction on site selection requested by SERDP was that most, if not all, of the sites should be located on military installations. This condition required, of course, that permission for the work from the installation commanders be secured through the installations' Cultural Resource Management (CRM) programs.

3.3.1. Site Selection Issues

The sites included in this project were selected through consideration of three key factors—contrast, clutter, and environment—as well as several secondary issues. Since these factors should be considered whenever one assesses the feasibility of investigating a particular site using geophysics or other forms of remote sensing, it is appropriate to discuss them here (see also Section 7 for theoretical perspectives).

Contrast is the single-most important determinant of the potential for detecting subsurface archaeological features using geophysical sensors (see Section 7 and Kvamme 2006:206,208; Somers 2006:112). Features are detectable if they exhibit sufficient contrast with their immediate surroundings in terms of at least one of several physical properties: magnetism, electrical conductivity, or dielectric variations (Weymouth 1986; Conyers 2006). Contrasts detectable from the air or space depend on other factors and include soil density, particle size, type, nutrients, moisture, and other factors that influence patterns of surface vegetation growth, soil color variations, and thermal properties. Contrast is, of course, a product of an archaeological feature's characteristics and its surroundings. The former includes a feature's material composition, dimensions, and volume. Since archaeological features typically exist as discrete entities within a matrix of soil, characteristics of soil (particle size, moisture, iron oxide content) help determine contrast as they differ from archaeological targets of interest.

Clutter is a characteristic of archaeological sites that diminish the potential for the detection of subsurface features by remote sensing methods (Kvamme 2006:222; Somers 2006:120). Clutter refers to variation in the data associated with a discrete source other than the phenomenon of interest. Characteristics of archaeological sites that often represent clutter include relatively shallow bedrock, rock inclusions in the soil, insect and animal burrows, tree roots, plow furrows, looters' holes, previous excavations, relatively recent metallic debris, and modern infrastructure (see Section 7). The basic objective of this project was to identify and, where possible, develop data processing methods for improving the potential of archaeological feature detection and interpretation. This goal could best be achieved using very high quality data from sites that contained a wide range of variation in archaeological feature types. Low signal-to-noise ratios and high levels of clutter limit the potential for identifying optimal approaches to data processing, data presentation, and particularly methods of data integration.

Site environment was a third factor that played an important role in site selection. Individual archaeological sites represent nearly unique combinations of environmental and cultural factors that contribute to contrast and clutter. The co-occurrence of these factors varies by region and creates complexities that make generalizations difficult. Although previous research has demonstrated that remote sensing can be used productively for archaeological detection throughout the United States, regional differences in environment as well as cultural forms often influence the choice of sensors and call for adjustments in survey strategies and data processing.

In the arid Southwest, for example, late prehistoric sites are often characterized by coarse soil with low moisture content. Pit houses later gave way to surface rooms constructed of native stone or adobe. Dry soil may preclude resistivity surveys but facilitate radar investigations; poorly developed soil may limit the usefulness of magnetometry; lack of vegetation means that vegetation marks detectable from the air may not be formed. In the Great Plains and Midwest, however, late prehistoric habitation sites occurred in loamy soils with greater moisture and ample surface vegetation, where semi-subterranean pit houses and substantial storage facilities widely occur, but stone and adobe were rarely used. In this region, resistivity, magnetic, and aerial surveys could be of great use, but radar penetration might be more difficult in the moist and conductive soils. The Southeast is characterized by relatively warm temperatures, abundant moisture and vegetation, and in many areas, soils rich in iron oxides. In much of this region late

prehistoric houses were characterized by daub wall cladding that is readily detected by magnetic methods.

Broad temporal trends in prehistoric adaptations are also relevant to the usefulness of remote sensing to the archaeologist. Throughout most of prehistory, people were hunters-and-gatherers who relied on residential mobility to adjust to seasonal and localized variation in resource availability. Early occupations are therefore rarely characterized by substantial constructed features that might be remotely detected (e.g., storage pits, houses, hearths, and graves). In most regions, populations became more sedentary only as they began to rely more heavily on agriculture. Later prehistoric habitation sites that were occupied on a multi-seasonal or even a year-round basis are far more likely to include substantial constructed features, such as storage facilities and houses as well as deposits of organic middens. Habitation sites dating to the late prehistoric period are not only very amenable to remote sensing, they are also typically viewed by cultural resources managers as being more likely to provide scientifically and culturally valuable information. Historic sites of the later nineteenth century are typically characterized by the use of construction materials that include ferrous metal nails and other fittings, metal pipes, brick and concrete. These materials usually create high-contrast anomalies in geophysical and other data that reduce the importance of variation in soil, moisture, and other site characteristics (Kvamme 2006:228).

3.3.2. Site Selection Goals

Given the foregoing considerations, the SERDP team decided to select archaeological sites for study that met the following conditions.

1) *Diverse environments*. As a group, the selected sites should occur in diverse environments that include distinct differences in soil characteristics, particularly particle size and moisture retention.

2) *Variation in archaeological features.* The sites should also reflect a wide range of variation in the types of archaeological features expected to be present. This was essential to allow the project to examine differences between sensors and combinations of sensors in terms of their ability to detect features of varying size, composition, depth, and contrast.

3) *A range of time periods*. Although it was necessary to select relatively late habitation sites to ensure the presence of diverse archaeological features, it was also desirable to achieve a reasonably wide range of variation in the time periods represented.

4) *Reasonably intact sites.* It was important to choose sites where one could expect subsurface features to be relatively well preserved. Sites that had been significantly impacted by heavy vehicle traffic, extensive looting, or large-scale archaeological excavations were viewed as poor candidates.

5) *Sites free of metallic debris.* The sites needed to be reasonably devoid of metallic clutter. Military installations are frequently characterized by a generalized scatter

of recent metallic items, much of it the result of decades of military training. While metallic debris poses only minor problems for the use of some remote sensing techniques, it significantly limits the potential for effective magnetic surveys, and impacts other geophysical sensors.

6) *Sites free of dense vegetation, trees, and survey impediments.* It was necessary to choose sites where the vegetation cover would not seriously complicate data collection. All of the near-surface geophysical techniques used by archaeologists require a sensor to be moved systematically across the site surface. Vegetation, rocks, steep ground that impeded such controlled movement would increase problems with noise and clutter and significantly reduce the rate of data collection. Tree cover would preclude aerial and space imaging.

7) *Sites of archaeological and historic interest.* A final but important consideration was a site's ability to arouse the interests of a broad range of archaeologists, historians, and resource managers. It was assumed that most individuals with an established interest in the use of remote sensing would be interested in the project's results. It was hoped that a much broader interest in this research could be ensured by working at locally or regionally important sites that would be of great interest to archaeologists who presently have little interest in or familiarity with these methods.

3.3.3. Candidate Sites

The effort to identify archaeological sites that met these criteria began with numerous informal requests for suggestions from CRM personnel at a number of Army installations, including Fort Benning, Fort Bliss, Fort Bragg, Fort Drum, Fort Hood, Fort Leonard Wood, and Fort Riley. Similar requests for suggestions about suitable sites were extended to colleagues working with CRM firms in the Southeast, Midwest, and West, as well as colleagues in the remote sensing and geophysics industries. A number of interesting sites were suggested, but many failed to meet one or more of the specified criteria. Some of those that did represent good candidates in most respects did not meet the criterion of "great intrinsic interest."

Sites located at Fort Benning, Fort Bliss, Fort Drum, and Fort Hood were initially viewed as serious candidates but were ultimately not selected. Notable among these was the Yuchi Town site (Hargrave et al.1998), a late prehistoric through seventeenth century occupation at Fort Benning. Excavations at Yuchi Town in 1994 and 1995 encountered well-preserved remains of substantial late prehistoric houses and other features including burials. Yuchi Town was eventually rejected for several reasons. Those portions of the site known to include diverse features were densely forested. The vegetation cover would have precluded the use of aerial and satellite imagery and significantly slowed the rate of survey using near-surface techniques. There was also a concern that the abundant looting at Yuchi Town would cause too much of a problem with clutter. Additionally, the site was in a relatively remote location (most easily reached by boat from the Chattahoochee River), so logistics and increased costs would have been a concern given the need to transport a considerable amount of equipment.

Fort Hood offered the use of several interesting sites. The best candidate, 41CV1141, was expected to include basin hearths, slab-lined pits, large mounds of burned rock, and middens. It was located in an area of shallow sandy soils covered in grass and sparse cedars. The site was ultimately not selected because of possible impacts from tank traffic and erosion.

At Fort Bliss, the sites of several adobe pueblos (Twelve Room Pueblo, Sgt. Doyle, and Pueblo Escondido) appeared to be excellent candidates and the largest of these, Pueblo Escondido, was included in this study. A nineteenth century site, the Turquoise Railroad site, was briefly considered. Although it included a number of wood and brick structures, it too was eventually rejected in favor of Pueblo Escondido.

Fort Drum, located in upstate New York, offered several interesting candidate sites, including FDP1093. This was an Early Woodland occupation (ca. AD 200) known to include hearths and storage pits. It was located in an area of sand dunes covered in low grass. Ultimately, FDP1093 was not selected because it was thought that other sites would offer a greater diversity in feature types.

3.3.4. Selected Project Sites

Four sites were ultimately selected for inclusion in this project: Army City (Fort Riley, Kansas), Pueblo Escondido (Fort Bliss, New Mexico), Kasita Town (Fort Benning, Georgia), and Silver Bluff (located near the DOE Savannah River facility, Aiken County, South Carolina) (Table 3.1). Each of these sites is described here, with a focus on information that was available prior to SERDP investigations. New information about the sites that resulted from the SERDP work is presented in subsequent sections.

Table 3.1. *Characteristics of sites selected for inclusion in the SERDP study.*

Site	Location	Dates	Soil	Architecture
Army City	Fort Riley, KS	1917-1921	Silt loam	Commercial buildings, dirt streets, concrete, iron, and wood construction
Silver Bluff (G. Galphin site)	Aiken Co., SC	1752-1780	Sand	Defensive stockade, residential buildings, brick and wood construction
Kasita Town	Fort Benning, GA	1725-1825	Sand, sandy loam	Traditional residential buildings, wood and daub construction, possible nails
Pueblo Escondido	Fort Bliss, NM	1250-1450	Silt loam	Traditional residential buildings, pit houses and adobe surface rooms

3.3.4.1. Army City, Kansas

Army City was a private business venture established in 1917 to provide goods and services to the thousands of soldiers being trained at nearby Camp Funston (Fort Riley), Kansas (Hargrave et al. 2002; Rion 1960). The Fort Riley region is characterized by a temperate continental climate. August is the warmest month, with daily highs exceeding 90° F, and daily lows that average about 66° F. January is the coldest month, with the daily average temperature ranging between 18° and 40°. Average yearly precipitation is 31.64 inches, occurring mostly as

rain (Jantz et al. 1975; Kresja 2005:7). The site is located on the floodplain of the Kansas River immediately beneath nearby bluffs. The uppermost soils consist of a silty loam that is underlain by a silty clay loam. Since its abandonment about 1921, the site has served as a hay field and its archaeological deposits lie just below the surface. The only surface indications of Army City are a few depressions in the ground, pieces of concrete footings, and (under favorable conditions) vegetation marks that show faint outlines of former structures and streets (see Section 5.12.3). Its archaeological deposits have not been impacted by plowing (probably because of the presence of occasional concrete or other rubble near the surface) and disruption by tree roots is almost non-existent. A recent QuickBird satellite image of Army City, of panchromatic data at a spatial resolution of 60 cm, reveals the hayfield in which the former town lies, the bluff to the north, the town of Ogden to the east, the study boundary over the town's commercial core used by this project, with the loci of several historic streets superimposed (Figure 3.1).

Most of the available information about the history and day-to-day operation of Army City is derived from a Master's thesis by George Rion (1960). He reviewed newspaper accounts and the few official records that survived a 1920 fire and the major flood of 1935, and interviewed local informants including one of town's founding fathers. Rion's thesis includes several photographs of Army City, and additional images have been secured from the Kansas State Historical Society. One of the most informative of these photographs is a panoramic image of Army City that was taken (probably in 1917) from atop the nearby bluffs (Figure 3.2). Additional photographs show the facades of selected buildings soon after their construction, soldiers and residents wading through ankle-deep water during a 1919 flood, and the 1920 fire.

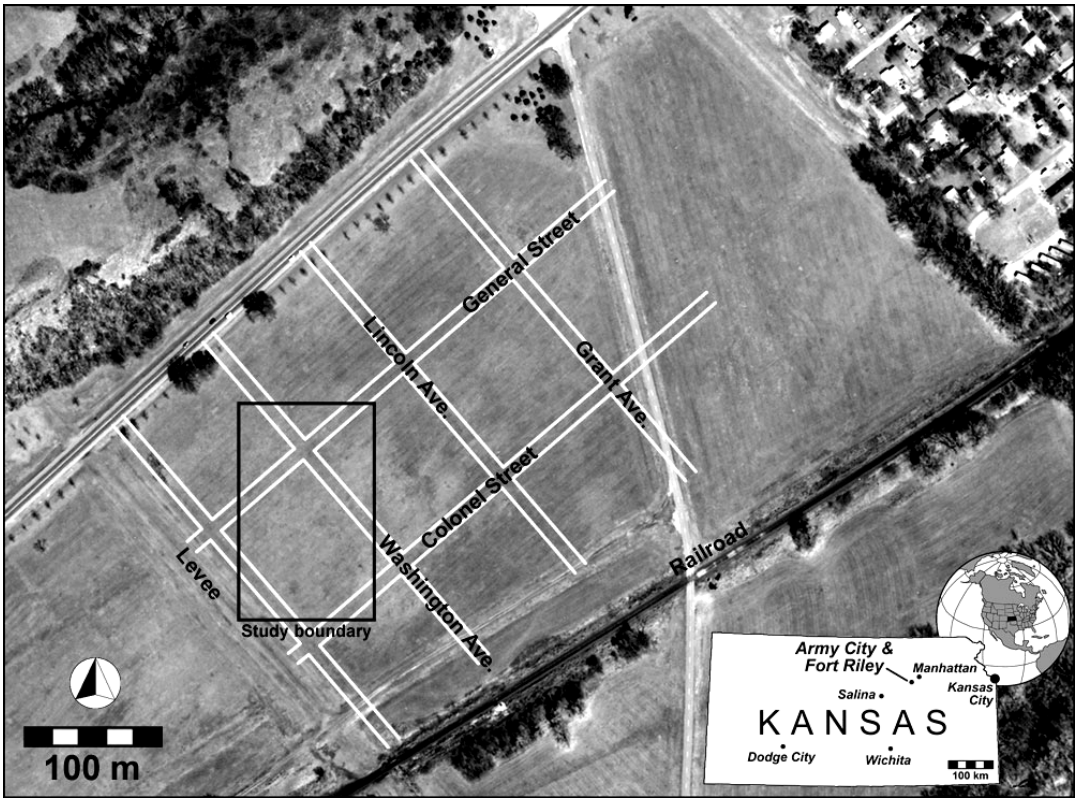


Figure 3.1. *Present-day Army City showing portions of its former street grid and the study area centered over the town's commercial district. Relatively few structures were built outside of the core area. (Image source: QuickBird panchromatic at 0.6 m nominal resolution, dated April 3, 2001. © Copyright 2002, DigitalGlobe Incorporated, Longmont, Colorado, USA, used under license agreement.)*

Figure 3.2. *Panoramic photograph of the commercial core of Army City shortly after its construction in 1917 (cropped from original panoramic view; photo credit: Kansas State Historical Society). View from bluff top looking southeast down Washington Street. The Hippodrome is the largest building in the center; the Orpheum Theatre lies across the street to the left; Camp Funston lies to the far right.*

Army City was a planned community in that all of its buildings were constructed over a very brief time interval, and all were required to conform to a "Spanish mission" architectural style. This was achieved largely through the use of stucco wall cladding and design features such as flat roofs. As platted, the town included five north-south and four east-west streets (Figure 3.1), although the southern-most area apparently was never developed. The town's plat indicates that the unpaved streets were 60 feet wide (although our remote sensing shows 35 feet widths; 60 feet only when sidewalks on both sides are included); utility pipes ran beneath the 16 feet-wide alleys (but remote sensing shows a more general distribution of pipes). Portions of the commercial district boasted electric street lights, concrete sidewalks, and many buildings had telephones. Army City's business district in the northwest portion of town included two large movie theaters (the Hippodrome and the Orpheum) that provided a combined seating of 1,500, as well as restaurants, barber shops, photography studios, civilian and military clothing stores, gift shops, a bank, garage, lumber yard, pool halls, and other businesses. Although it was, in many ways, a lively "pay-day town," by urban standards Army City was a relatively tame place. Only low-alcohol content beer was sold and efforts to prevent prostitution were reportedly successful (Rion 1960). Like most aspects of American society at the time, the army was racially segregated. A separate complex of buildings located just south of the railroad track that bisected Army City housed businesses that served the African-American soldiers (Figure 3.2). The eastern portion of the town north of the railroad was a less densely developed mix of modest private residences, businesses (a dry cleaners and a hotel), and small public buildings (including a Lutheran center) (Hargrave et al. 2002; Rion 1960).

Despite occasional interruptions to commerce that resulted from military training schedules and epidemics that included the influenza outbreak of 1918, Army City flourished until the end of the war. After the armistice was signed, many soldiers came through Fort Riley to be discharged, but it appears that few remained long enough to benefit the town. In August of 1920 a fire devastated several blocks in the commercial district and this appears to have signaled the town's inevitable demise. Most businesses ceased operation soon after the fire. Buildings not destroyed were either auctioned and dismantled or, in one or two cases, moved intact to the nearby town of Ogden. As a result of a land purchase by the government during the World War II era, the site of Army City is now located within the bounds of Fort Riley (Hargrave et al. 2002; Rion 1960).

In 1996 and 1997, CERL assisted Fort Riley in evaluating Army City's eligibility for the National Register of Historic Places (NRHP). Although Rion's (1960) study made a strong case for the site's historic significance, almost nothing was know about the quality of preservation of the archaeological deposits. Given the site's size, a thorough investigation using hand excavation would clearly have been an expensive proposition. It therefore was decided to investigate the site using geophysical surveys and small-scale but carefully targeted excavations. Dr. Lewis Somers of Geoscan Research USA conducted a small-scale geophysical survey in the eastern residential area in 1996 using electrical resistivity, magnetic field gradiometry, and ground penetrating radar. While the resistivity and gradiometry data both produced useful information, the former received focus. In 1997 Geoscan Research USA conducted a low-density (1 reading per square meter) resistivity survey that covered approximately 9.2 ha. This represented most of the site north of the railroad, and roughly one-half of the total site area estimated by the

town's plat (although most of that area was never filled in with constructions owing to its short life). That effort was quite successful, yielding a rather dramatic, highly detailed map showing the location of streets, alleys, sidewalks, utility trenches and pipes, and massive deposits of architectural debris (Figure 3.3). This data set was extremely important to the project because it showed exactly where subsurface features might be found within the site and helped guide placement of our study area (Figure 3.1). Excavations were conducted in 1996 to validate indicated anomalies and more concretely identify the nature and character of their sources. This work focused on the eastern portion of the site and consisted largely of "shovel test pits" (small 50 x 50 cm excavations) and several larger excavation units a few meters in size. The following year, small-scale excavations in the commercial district documented the massive character of deposits of architectural debris. The site was determined to be eligible for nomination to the NRHP, and the combination of large-area geophysical surveys and targeted "ground-truthing" by excavation was recognized as an effective approach for investigating large complex sites (Hargrave 1999a, 1999b; Hargrave et al. 2002; Kreisa and Walz 1997; Larson and Penney 1998; Somers 1997, 1998).

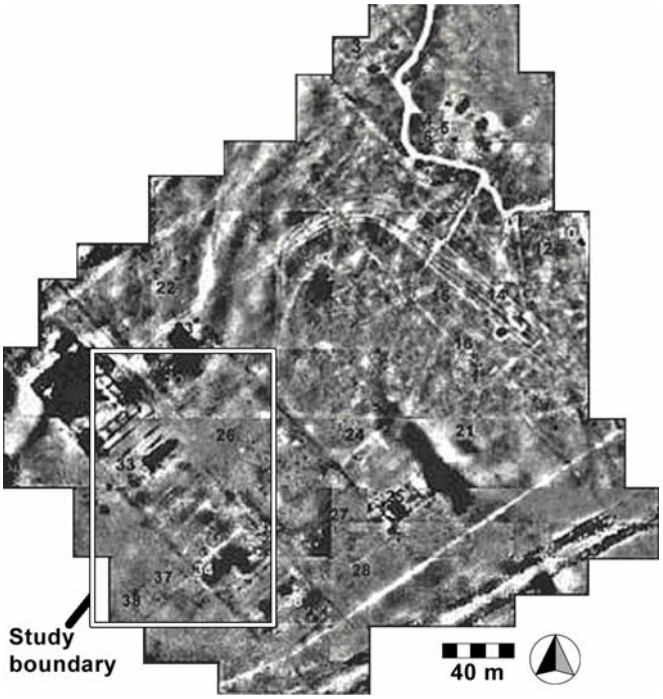


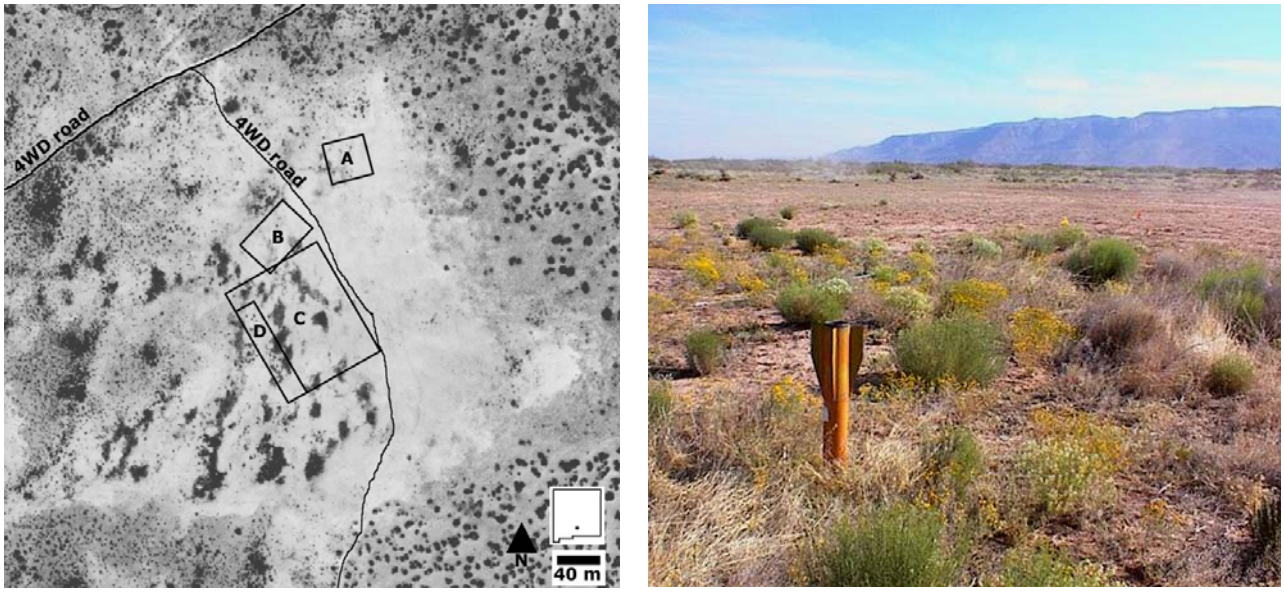
Figure 3.3. *Electrical resistivity survey of Army City conducted in 1997 by Geoscan Research USA showing that most buried cultural features exist near the town's commercial core to the west. The high bluffs lie to the upper left and the railroad tracks in the lower right, immediately off the image area. The SERDP Project's study boundary is superimposed.*

On balance, Army City met most of the criteria for inclusion in this project and, as will be seen in subsequent chapters, produced very high quality geophysical data. As expected based on its early 20th century date, the presence of abundant ferrous metal, concrete and masonry building materials resulted in very high contrast anomalies. While the geological background contributed some clutter, it did not pose a serious problem. Army City is perhaps the only civilian-owned planned community designed to serve U.S. soldiers during the First World War. As a segregated town, Army City appears to offer an

excellent opportunity to study the material correlates of racial biases in access to goods and services. It is reasonably well-preserved, although the archaeological deposits are dominated by massive amounts of building materials and a relatively low density of artifacts related to food processing, consumption, and other activities. Given these characteristics, Army City is a site of great historical interest and significance and was an excellent choice for inclusion in this project.

3.3.4.2 Pueblo Escondido, New Mexico
Michael L. Hargrave, Construction Engineers Research Laboratories, and Eileen G. Ernenwein, University of Arkansas

Located approximately 20 miles south of Alamagordo, in Otero County, New Mexico, Pueblo Escondido is jointly managed by Fort Bliss and the Bureau of Land Management (Figure 3.4a).



a **b**

Figure 3.4. *The site of Pueblo Escondido, New Mexico. a) Satellite image showing patches of native vegetation, access roads, and loci of the four project study blocks, with C-D forming the principal study area (image source: QuickBird panchromatic at 0.6 m nominal resolution, dated March 26, 2003. © Copyright 2000, DigitalGlobe Incorporated, Longmont, Colorado, USA, used under license agreement). b) Surface view showing vegetation cover and imbedded missile that served as a site datum.*

The site locale occurs within the Chihuahuan biotic province, characterized by very low levels of precipitation (the yearly average is about 10 inches), high rates of evaporation, and a general absence of surface water. Most precipitation occurs in the form of summer thunderstorms. Temperature is lowest in December and January, with freezing temperatures common from late November through early March. June, July, and August are the hottest months, with daily highs often exceeding 100° F (Seaman et al. 1988:24). The site is situated on an alluvial plain surrounded by sand dunes (Hedrick 1967). Sediments within the site are massive silty loams with few alluvial gravels (TRC 2005:35). Water sources include seasonal run-off from the Sacramento Mountains via a network of shifting arroyos (Anyon 1985:8). A substantial dry stream is located about 1

km south of the site and was probably closer in the past. The site has sparse vegetation, although there are a number of relatively dense stands of brush that includes yucca, rabbit brush, palo verde, snakeweed, mesquite, and various grasses (Figure 3.4b).

Although professional archaeologists have been aware of Pueblo Escondido since the 1930s, it was first officially documented in May 1953 by the El Paso Archaeological Society (EPAS Site Survey Form). Since then, archaeologists have visited the site on a number of occasions. Intense looting and shifting sand dunes along the site margins have made it difficult to ascertain basic issues such as site size and condition. It is likely that earlier professional observations can provide more information about some aspects of the site (for example, site lay-out and minimal number of rooms) than more recent observations. In 1968, John Green and John Hedrick estimated site size at .3-mile (ca. 483 meters) north-south by .2 mile (ca. 322 m) east-west. Four parallel, east-west oriented rows of pueblo-style rooms were visible, and the site was believed to include 60 to 80 rooms. Overall site condition was characterized as good, although there was evidence that six rooms had been vandalized. The character of the room fill was described as "fallen adobe wall material, sand and lake silt." Fragments of burned posts and roof beams were observed in the vandalized areas. Room floors were found to be 30 to 60 cm below present ground surface. Artifacts recovered at the site included triangular and notched projectile points, disc and olivelle beads, turquoise and shell fragments, and large quantities of lithic and ceramic artifacts (Hedrick 1968).

Investigations of the site conducted by the El Paso Archaeological Society (EPAS) in 1965 resulted in a somewhat different estimate of site size (364 meters east-west by 113 meters north-south), and site depth was estimated at 10 to 20 cm (Hedrick 1967:19). The EPAS investigations included a controlled surface collection of 10 percent of an 80 by 60 m area (Hedrick 1967:19). Six 1 by 1 m test units were excavated within this area, but only one of these produced cultural material. Additional excavations consisting of a two 1-m wide trenches (arranged so as to form a T) investigated one of two contiguous rooms.

Room 1 appears to have been nearly square, measuring about 5.7 m on a side. Remains of the walls and floor were plastered adobe; with an average wall thickness of 34 cm. Internal features included a rectangular adobe step, a plastered adobe hearth, four substantial postholes, and evidence for secondary posts. Fragments of roof beams were present but did not provide clear evidence for roof construction techniques. The ceramics (primarily El Paso Polychrome) and architectural characteristics indicated that the site dated to what is termed the El Paso phase (A.D. 1200-1400) of the Jornada Branch of the Mogollon culture (Hedrick 1967:20; Lehmer 1948).

Pueblo Escondido was again documented in 1975. At that time site condition was described as poor to fair. The site form notes that "In the central area of the site, the position of many structural remnants appear to be indicated by debris concentrations resting on level or slightly elevated mounds…Although the site has been extensively vandalized and has been exposed to vehicular traffic, much of the site area appears intact" (Beckes 1975).

Test excavations conducted in 1975 by the Texas Archaeological Survey revealed that Feature 1, a large, rectangular surface concentration of artifacts was associated with the remains of a hard-packed adobe floor at 30 cm below ground surface. Dimensions of the structure were found to be 9.4 by 8.4 m. On three sides, the structure walls were about

30 cm high and 35 cm thick. The eastern wall was thinner (ca. 20 cm), suggesting that it was an interior wall that separated the investigated room from a neighboring room in the same room block (Beckes 1975). Internal features associated with the investigated room included a looted burial pit, a possible posthole, an irregular depression, a circular pot rest depression, and a rectangular depression that was associated with a displaced diorite trough metate found nearby in looter's back dirt. Small concentrations of carbonized wood found on the structure floor appeared to be the remains of roof members or cross-beams. Diagnostic ceramics again suggested that most of the occupation at Pueblo Escondido dates to the El Paso phase, A.D. 1200-1400. A radiocarbon assay on a sample of in-situ roof elements yielded a date of A.D. 1260 +/- 70 (TX2341), and this date is fully compatible with the El Paso phase's date range of A.D. 1200-1400 (Beckes 1975).

Archaeologists classify the culture at Pueblo Escondido as *Mogollon*, which is one of the three major cultural traditions in the American Southwest. It is the largest in area and subdivided into several branches, including the Jornada branch where Pueblo Escondido resides (Lehmer 1948). In the Jornada region, pithouse villages are documented by A.D. 400 or earlier, and by approximately A.D. 1200 the first pueblos appeared (O'Laughlin 2001). These changes occurred within what is termed a Formative Period in North American archaeology, which marks a major change in lifestyle from hunting-and-gathering to one that includes the use of pottery and dramatic increases in population, sedentism, reliance on domesticates, and cultural complexity. In the Mogollon area it is thought that seasonal mobility and the use of wild plants and game remained an important strategy and that completely sedentary living was never practiced (Whalen 1994). Nevertheless, extensive pithouse villages and eventually pueblos were built, suggesting sedentary living for part of the year. The El Paso phase (A.D. 1200-1450) represents a high point in terms of population levels, sociopolitical complexity, ceremonialism, and inter-regional interaction with the appearance of above ground adobe pueblos comprised of contiguous rooms (TRC 2005:23). El Paso phase sites are known to exhibit two community plans. Smaller settlements were comprised of linear (generally east-west oriented) complexes of contiguous rooms. In larger settlements, room blocks (sometimes involving more than 100 rooms) were distributed around a central plaza. El Paso phase pueblos were typically single-story and sites ranged in size up to about 10 hectares. Pit house structures occur throughout the phase, although they appear to have assumed a secondary role. At nearby Firecracker Pueblo, non-contiguous pit houses arranged in linear patterns date to the early El Paso phase and are followed by the appearance of pueblo construction (O'Laughlin and Martin 1993:35; TRC 2005:24). Pueblos are primarily found near playa lakes and alluvial fan margins where water allowed horticultural subsistence. Hunting activities were primarily focused on small game including rabbits (Whalen 1994). There is an increase in the size and density of habitation sites, suggesting a population increase and higher levels of social organization.

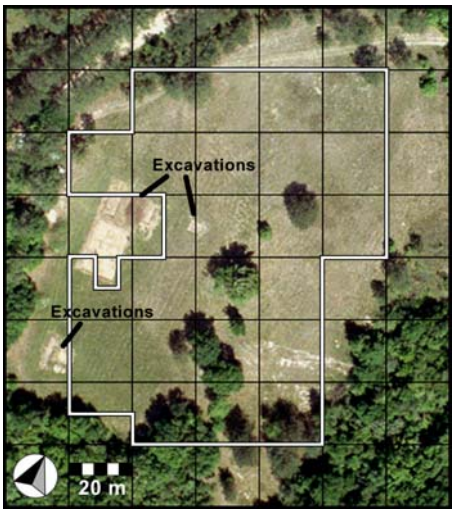
Many of the changes observed in architectural styles and religious iconography suggest a strong influence from Paquimé (Casas Grandes) in northern Mexico, which began expanding its sphere of influence around A.D. 1200 (Miller and Kenmotsu 2004). Mauldin (1986) suggests that large, more permanent sites situated along alluvial fan margins organized logistical use of adjacent resource areas for hunting, foraging, and horticulture maintenance. During the summer much of the population left the large pueblos and dispersed in smaller settlements where hunting and gathering was more

productive. Aggregations at large pueblos during fall for harvest and spring for new planting would have been necessary. The El Paso Phase represents a climax in population, horticulture, and cultural complexity in the region, which was followed by a collapse and apparent abandonment of the area or dramatic shift in lifeways by A.D. 1450. Reasons for abandonment have long been debated. They may include a combination of overpopulation, agricultural devastation, and drought that made occupation of the region non-viable.

Overall, Pueblo Escondido was an excellent site for this project. It has long been recognized as the largest settlement in the area, yet only very limited investigations have been conducted. Basic issues about the site, such as the number and arrangement of rooms, remained unknown, due in part to the presence of sand dunes around the site margins. Previous excavations had documented the presence of intact architectural remains. Despite these factors, there was initially some uncertainty as to the extent to which adobe features would contrast with surrounding soils, and the extent to which reported looting would cause excessive clutter. Fortunately, site conditions at Pueblo Escondido proved to be very favorable to geophysical surveys.

3.3.4.3. Silver Bluff Plantation, South Carolina (The George Galphin Site)

Silver Bluff Plantation (or more simply Silver Bluff) is a name often used to describe the George Galphin site (38AK7). It is located in Aiken County, South Carolina, and was placed on the NRHP in 1977. This 128-acre property includes the archaeological remains of a trading post founded by George Galphin in the 1750s that played a central role in the historic settlement of the region (Herron and Moon 2005).



a

b

Figure 3.5. *The Silver Bluff site. a) Aerial view showing loci of archaeological excavations and geophysical survey boundary. The Savannah River lies immediately to the south. b) Ground view to the south. The large excavated area lies under the white tarp.*

Silver Bluff is located in the Atlantic Upper Coastal Plain, in an ecotonal area between the oak-hickory and magnolia forests of the Southern Temperate Deciduous Forest Biome (Shelford 1963). The site is flat and slopes very gently down to a steep

bluff edge overlooking the Savannah River (Figure 3.5). The site's soils are classified as Troup Sand (0-6 percent slope), which has a low organic content, drains rapidly, and is now generally planted in loblolly and longleaf pine (Herron and Moon 2005; Rogers 1985). Although it is not well-suited for agriculture, the site was cultivated from Galphin's time to the early 1980s. Galphin grew corn, tobacco, and indigo, whereas corn, cowpeas, soybeans, and peanuts were cultivated during the late 19th and 20th centuries (Herron and Moon 2005).

George Galphin arrived in the colonies from Ireland in the 1730s and soon established himself in the deerskin trade. As a backcountry trader he was trusted by both colonials and natives and served as an interpreter in the colony's interactions with the Creek. He was a partner in the Augusta Company but established his own trading post at Silver Bluff and later founded a second post further west at Old Town. Galphin's will indicates that he constructed at least two brick buildings at Silver Bluff. During the 1760s and 1770s Silver Bluff gradually evolved from a trading post into a plantation. By the 1770s it served as a regional warehouse for tobacco awaiting government inspection. During the revolutionary war Galphin worked to maintain the neutrality of the Creek. Shortly after his death in 1780, Silver Bluff was occupied by British forces, but was recaptured the following year by South Carolina militia commanded by Colonel Harry Lee. Galphin's heirs sold Silver Bluff in 1792, and the property changed ownership several times during the early 19th century. It later became a productive but non-residential portion of a large, successful plantation, and the structures built by Galphin were used for storage. Silver Bluff was farmed by tenants in the years following the Civil War. The Galphin site is now located within the 3,154-acre Silver Bluff Audubon Center and Sanctuary (Herron and Moon 2005; Scurry et al. 1980).

The South Carolina Institute for Archaeology and Anthropology (SCIAA) conducted the first archaeological investigations at the Galphin site in 1979 and 1980. Surface collection of a systematic sample of the site area recovered approximately 9,000 artifacts. Analysis of the clay pipes resulted in an estimated date of 1761 whereas the ceramics yielded an estimated mean date of 1765. In 1996 the Savannah River Archaeological Research Program (SRARP) undertook investigations at the site that included a 100% surface collection of a 100 by 100 m area, a grid of closely spaced "shovel test pits" (STP; excavations about 30 x 30 cm in size), larger hand excavated units, and a ground penetrating radar survey; (results of the GPR survey were apparently inconclusive). In 1999 an archaeological field school was conducted at the site, sponsored by SRARP, Augusta State College, and the Audubon Society. Volunteers under the direction of Tammy Forehand Herron and Robert Moon (SRARP) continued small-scale investigations until 2002, when excavation ceased in anticipation of the SERDP geophysical survey. To date, 543 STPs have been excavated, 87 percent of which produced historic artifacts. Hand excavated units (largely confined to several excavation blocks) have exposed an area totaling 284 m^2 (Herron and Moon 2005).

The remains of nine buildings and a substantial stockade have thus far been documented. At least three of the buildings appear to be ephemeral structures that may have been the residences of slaves. A large brick building is identified as Galphin's residence. Another structure is, based on the artifacts recovered, likely to be the trading post's primary commercial building. Overall, the nine buildings are distributed in a manner consistent with a rectangular stockade. A section of a post and paling wall with

substantial postholes that supported large buttress posts was identified just west of the primary residence, and portions of the north stockade have also been identified. Research goals for the future include developing a better understanding of the nature and function of the remaining structures, identifying activity areas, and verifying the location of the west and south stockades (Herron and Moon 2005).

Silver Bluff clearly met many of the criteria for inclusion in the SERDP project. Both the site and its first owner played important roles in the early history of South Carolina and adjacent portions of Georgia, particularly in terms of relationships between the colony and native groups. Galphin's efforts to ensure the neutrality of the Creek nation were a significant contribution to the colonies' bid for independence from Britain. Thus, the Galphin site is clearly of great historic interest. The site's contribution to this project was, to some extent, diminished by the previous archaeological investigations, although this was not known to the SERDP team until the geophysical fieldwork was underway. Although a substantial amount of hand excavation had already occurred, most was concentrated within several contiguous excavation blocks. A more serious problem with clutter resulted from the extensive use of metal pin flags to mark the locations of the regularly spaced STP. Some of the pin flags rusted and broke into pieces too small to recover, others were cut and dispersed during a mowing of the site (Section 4.6.3), and some were inevitably lost. Unfortunately, these pin flags caused prominent anomalies in the magnetic data. While this metallic clutter does not preclude detection of the larger, higher-contrast anomalies associated with 18th century features, it may obscure anomalies associated with smaller, lower contrast features, and it greatly complicates the interpretation of ferrous metal artifact scatters. Anomalies associated with the pin flags also diminish the visual impact of the magnetic map, and this consideration is relevant in terms of efforts to impress a wide range of archaeologists with the usefulness and interpretability of geophysical imagery. This situation was made worse because, as it turned out, the magnetometry map was one of only three remotely sensed data sources procured at the site that yielded subsurface indications. These observations are not intended as a criticism of the Silver Bluff researchers, since metal pin flags are lamentably used by nearly all archaeologists in the U.S despite the availability of PVC pin flags.

3.3.4.4. Kasita Town, Georgia

Also known as Cussetuh (along with several other spellings), Kasita Town (9CE1) is located at Lawson Air Field, Fort Benning, Chattahoochee County, in extreme western Georgia. Fort Benning occurs at the extreme northern margin of the Coastal Plain physiographic province (Fenneman 1938), and within the Southeastern Evergreen Forest Region defined by Braun (1950). Kroeber (1963) describes the modern flora as mesophytic evergreen forest with pockets of swamp land. The region is characterized by long, hot, humid summers and mild winters. Annual precipitation averages 48 inches, nearly all of it in rainfall, with only traces of snowfall (U.S. Army Engineer Topographic Laboratories 1976:27). The Kasita Town site is situated on a level terrace within 200 meters of a great bend in the Chattahoochee River. A recent soil map classifies the SERDP survey area (as well as most of Lawson Air Field) as "Udorthents-Urban Land complex, 0-10 percent slopes" (O'Steen 1997:10-11), which is clearly not particularly

useful to the archaeologist. The 1928 soils map indicates that the study area lies within two soil types: Cahaba Sand and Cahaba Sandy Loam (2005b:32).

A review of the Georgia site files (O'Steen et al. 1997:22) revealed that there has long been some uncertainty as to the location of Kasita Town. The site's history is complicated by the fact that the town was abandoned and then re-established at least once. An earlier settlement, known as Old Kasita Town, is believed to have been located on higher ground within the main post. It was probably established prior to 1685 by Native emigrants from the Alabama and Coosa River valleys (O'Steen 1997:396). The settlement investigated in the SERDP project is almost certainly the latest Kasita Town, established in the mid-18th century, when the Creek returned to the Chattahoochee area following the Yamassee war (ca. 1715). The site was probably occupied until the Creek were forcibly removed from the area in 1825. Beginning in the mid-20th century, portions of the site have been impacted by infrastructure associated with Lawson Air Field, including runways, massive drainage pipes, and a road and fence that (on the west side of the site) parallel the river (Foster 2005a, 2005b; Figure 3.6).

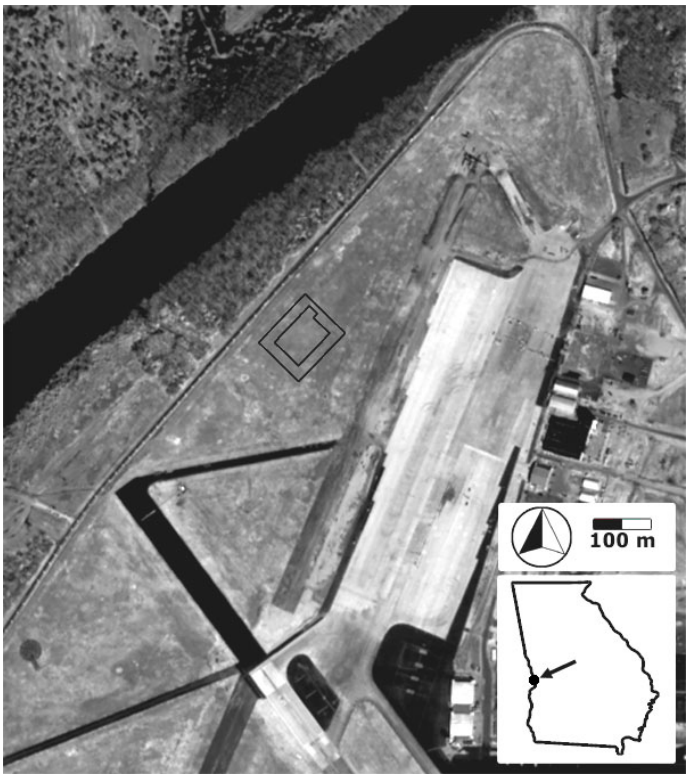


Figure 3.6. *Satellite view of the SERDP Project study area immediately west of Lawson airfield and above the Chattahoochee River (Image source: QuickBird panchromatic at 0.6 m nominal resolution, dated March 13, 2003. © Copyright 2000, DigitalGlobe Incorporated, Longmont, Colorado, USA, used under license agreement.)*

Modern studies of the historic Creek Indians benefit from first-hand accounts recorded by William Bartram, David Taitt, Benjamin Hawkins, and others (Foster 2003, 2005). Based on such sources, it is known that Creek towns conformed to a traditional layout. At a town's center was a square ground, a "chunky yard," and the circular winter council house (also referred to as a rotunda or "hot house"). The square ground was bordered on each side by a rectangular structure. These centrally located structures and facilities were surrounded by domestic compounds, each consisting of a cluster of structures and small garden plots (Foster 2005; Hudson 1976:214; Swanton 1979;

Waselkov and Braund 1995). The Kasita site is reported to have been bisected by another historically important feature, the Federal Road. Upgraded from a horse path in 1811, the Federal Road was suitable for wagon traffic, and by 1820 was an established mail and stagecoach route (O'Steen et al. 1997:90-91).

Several archaeological investigations were conducted at the Kasita Town site prior to the SERDP study (Cottier 1977; Willey 1938). In 1936, Gordon Wiley excavated a number of test units to investigate the site prior to the construction of Lawson Field. Willey identified Creek structures and burials that contained 18th century trade items (Foster 2005b; Willey and Sears 1952). The exact locations of Willey's units and the Federal Road remain uncertain, although Foster (2005) believes they were located just outside of the SERDP geophysical survey area. The next major investigation of Kasita occurred in the context of a cultural resource assessment of 4,690 acres conducted in 1996 by New South Associates (O'Steen et al. 1997). Within the airfield, shovel tests were excavated at 15-meter intervals, providing an unparalleled opportunity to map the distributions of various artifact categories. Approximately 4,000 diagnostic sherds were recovered in the general vicinity of site 9CE1 (interpreted as Kasita Town) (O'Steen 1997:347). The New South investigations demonstrated that the site is, in many ways, relatively well-preserved, despite various impacts stemming from the construction and expansion of facilities associated with the airfield. Based on ceramic evidence, it was possible to identify the spatial extent of three occupational components associated with the Blackmon Phase and early and late Lawson Field phases (O'Steen 1997:383). For each of these components, artifact distributions were used as a basis for postulating the locations of residential areas as well as the "community ground" (i.e., the town's centrally located public structures and facilities) (O'Steen 1997:384-389).

In 2001, Panamerican Consultants undertook NRHP evaluations of portions of the 9CE1 site that would be impacted by an expansion of the airfield facilities. That work, located approximately 300 m east of the SERDP survey area, evolved into a large-scale data recovery project that included magnetometry survey. Results of the total field magnetic survey appear to have been seriously compromised by the use of widely separated (5 m) data collection transects and the abundant metallic infrastructure near the existing runway (Geoarchaeology Research Associates 2000). The subsequent data recovery effort included the hand excavation of 320 m^2 and the mechanized stripping of 45,000 m^2. The 348 excavated features included postholes, "trash" pits, "smudge" pits, and five burials (Foster 2005b).

Kasita Town is clearly a site of great historical and cultural significance and was, in that sense, a reasonable choice for inclusion in the SERDP study. Excavations that attempted to ground-truth or validate remote sensing findings (described in Section 5.16) revealed, however, that much of the study area had been impacted by grading associated with construction of the airfield. In some areas the A-horizon had been entirely removed; elsewhere it was preserved under a relatively thin layer of spoil. Unfortunately, these impacts do not appear to have been noted during the 1996 shovel test pit surveys. Roughly one-half of the 3,300 screened shovel tests produced artifacts, but they probably came from remnant bottom portions of cultural deposits below the zone of truncation. Nevertheless, artifact distributions provided a basis for inferences about temporal shifts in the site's spatial organization. The discontinuous areas of cutting and filling within the survey area represented clutter that diminished the potential for detecting anomalies

associated with subsurface features. In short, an awareness of such impacts is probably more important for those contemplating geophysical surveys than for those planning traditional excavations. In retrospect, the site should have been examined with a trench excavation in order to evaluate stratigraphy and depositional integrity prior to selecting Kasita Town for inclusion in the study. The remote sensing data from Kasita Town did not much contribute to the identification of sensors and processing methods that most enhance the potential for detecting subsurface features. The SERDP work at Kasita Town did, however, demonstrate the extent to which surface impacts such as grading can diminish the usefulness of geophysical surveys. In short, not all sites are good candidates for remote sensing investigations, and the ability to recognize poor candidates is just as important as the ability to identify ideal sites.

4. MATERIALS AND METHODS

This SERDP project employs ground-based geophysical techniques together with aerial and space-based imaging methods. The overarching purpose is to explore the benefits of multiple sensor modalities for detecting subsurface archaeological content and structure. This is accomplished through application to a variety of archaeological sites in several environmental settings. This section examines appropriate theory associated with each method, instrumentation, field and analysis methods, and initial results at each project site.

4.1. GEOPHYSICAL THEORY

Kenneth L. Kvamme, University of Arkansas

Geophysics is a complex topic that demands knowledge of certain fundamental principles in order that results can be fully appreciated. The following sections overview these principles.

4.1.1. Active and Passive Methods

Archaeologists utilize a variety of geophysical methods to measure physical properties of near-surface deposits—usually the uppermost 1-2 m, although deeper prospecting is sometimes practiced. The various methods each possess depth limitations that vary with particular soil properties and the nature of archaeological features being sought. Geophysical surveys in geology typically focus on *vertical* changes in physical properties through the subsurface (with common focus on stratigraphic sequences and deposits), often to great depths (Mussett and Khan 2000). Archaeo-geophysical surveys generally concentrate on *lateral* changes in the extreme near-surface in order to locate and define features of possible cultural origin (Clark 2000; Gaffney and Gater 2003).

Active and *passive* technologies are employed. The former might transmit an electrical current or radar beam into the earth to record responses; the latter measures inherent or native properties of the soil. Many geophysical techniques have been developed, but four are principally employed in archaeology owing to consistently good results (Gaffney and Gater 2003; Kvamme 2001): magnetometry, electrical resistivity, electromagnetic induction (EM), and ground-penetrating radar (GPR). All are active methods except the first. All generally respond to different and mostly independent dimensions of archaeological deposits (Weymouth 1986:371), but EM instruments mimic capabilities of certain other sensors.

4.1.2. Anomalies

This study is all about anomalies (first mentioned in Section 2.2), their definition, and interpretation. Useful results in geophysics are obtained when (1) archaeological features possess physical or chemical properties different from the surrounding matrix, and (2) instrumentation capable of measuring those properties is utilized with sufficient precision and sampling density to detect a contrast against the natural background in terms of magnetic, electrical, thermal, or other properties. These contrasts, referred to as anomalies, arise from anthropogenic causes, the targets of archaeo-geophysics, but also from geological, pedological, and biological phenomena (e.g., paleo-channels, tree

throws, animal dens). Sources of anomalies may be identified through excavation, deductive reasoning, or other means, topics that receive considerable focus below.

To understand many data integration methods, the character of anomalous measurements must be realized. By definition, they are extreme in value and differ from "normal" data from undisturbed deposits. Those larger than typical background values are referred to as "positive," while unusually small measurements are termed "negative." These concepts may be understood in terms of a statistical distribution where positive and negative anomalies exist on the right and left tails, respectively, while common background values define the central tendency (Figure 4.1). Robust anomalies include measurements of most extreme magnitude, located in far portions of the distribution's tails, relative to more subtly expressed anomalies lying closer to the center. In geophysics, unlike many applications areas, focus is placed on one or both *tails* of data distributions, a fact that profoundly influences analysis and methods of integration.

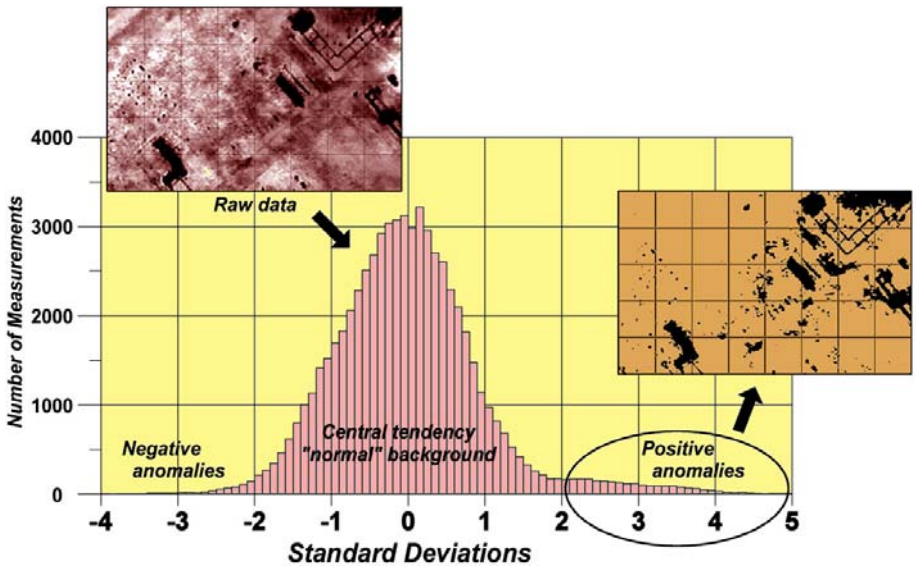


Figure 4.1. *Two-dimensional mapping of an electrical resistivity data set from Army City, Kansas (inset, top), a SERDP project site, and its one-dimensional histogram (after a normalizing power transform and standardization of the data; see Section 5.2.2). Normal background measurements occupy the central tendency while values in the tails are anomalous, indicated by a mapping of positive anomalies beyond two standard deviations of the mean (inset, right).*

4.1.3. Area Surveys and Sampling Density

Geophysical applications in archaeology commonly focus on lateral variations in the horizontal plane, which facilitates recognition of culturally generated patterns (see Section 2.2.2), making *area* surveys the norm in archaeology. Area surveys uniformly sample geophysical measurements over regions and the density, or spatial resolution, of those measurements in the *x-y* plane is a significant issue (see Clark 2000:158-164). Sampling density determines the size of archaeological features that can be resolved (Figure 4.2). A rule of thumb suggests the interval between data points should be no greater than *half* the size of the smallest feature to be detected (allowing multiple measurements to delineate it). Higher sampling densities potentially allow greater detail,

but there are practical limits. One is simply time. Instruments must be passed over the entire surface under investigation. More measurements per unit area mean more passes of an instrument requiring significantly more time. Twenty passes of a magnetic gradiometer were required in Figure 4.2a (1 m transect separation); 40 passes were required in Figure 4.2b (.5 m transect separation); 80 passes were necessary in Figure 4.2c (.25 m transect separation). The need for high resolution (one project goal) must be balanced against the need for surveying large areas (another project goal)—should an area be surveyed at *double* the normal density to better resolve small anomalies, or should *twice* the area be investigated? Related to this is the speed of instrumentation. An electrical resistance meter requires several seconds to acquire a *single* measurement, but ground-penetrating radar might acquire *50 or more* traces of data in a single second. Instrument speed obviously imposes limits on survey areas and data densities.

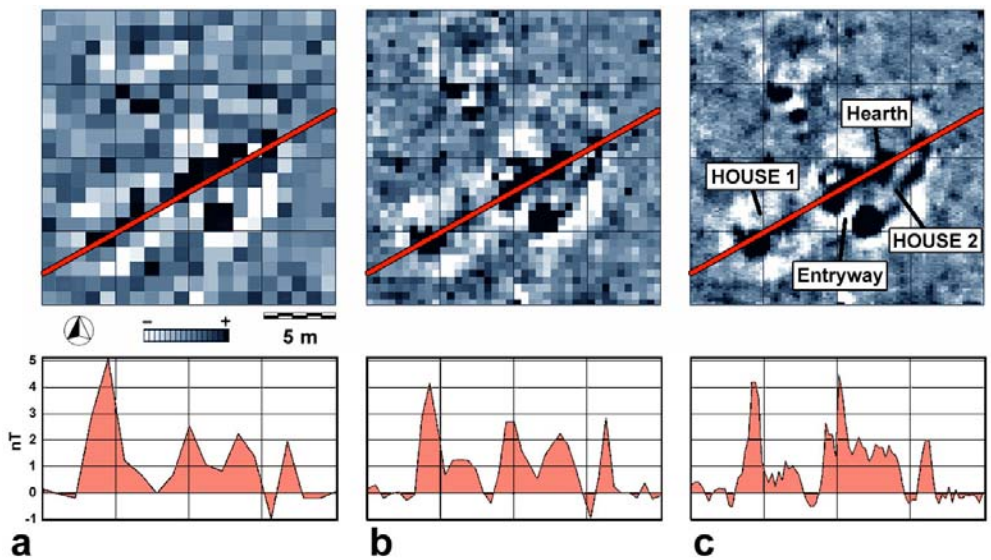


Figure 4.2. *Increasing sampling densities enable smaller anomalies and greater spatial detail to be detected, but significantly more time is necessary to acquire the data. Illustrated are partially burned and fully burned prehistoric houses revealed magnetically (after Kvamme 2003): a) 1 x 1 m samples, b) 50 x 50 cm samples, c) 25 x 12.5 cm samples.*

4.1.4. Magnetometry

Magnetometry is a passive detection method that measures the *sum* of remanent and all forms of induced magnetism below the instrument (whether natural or anthropogenic). Magnetic gradiometry is a form of magnetometry that records differences between two sensors, usually separated vertically (the differencing removes constant temporal changes in earth's primary field). Magnetic field strength is measured in *nanoteslas* (nT; 10^{-9} Tesla). In North America and much of Europe the background magnetic field strength ranges from about 40,000 to 60,000 nT (Weymouth 1986:341). This is noteworthy because magnetic anomalies of potential archaeological interest often lie well within +/-5 nT, and soil unit differences can be as subtle as 0.1 nT and less (recent work by Becker 1995, shows anomalies in the picotesla [.001 nT] range).

Magnetic survey instrumentation therefore is incredibly sensitive, capable of detecting less than one part in a half million.

Magnetometry has proven to be one of the most productive prospecting methods in archaeology for several reasons (Kvamme 2006). Intense heating of the soil beyond the Curie point (about 600° C) generates pronounced *thermoremanent* anomalies that typically arise from hearths, kilns, or the occasional burned house (Gaffney and Gater 2003). Magnetometry also responds to accumulations of magnetically susceptible materials (*magnetic susceptibility* refers to the ability of a substance to become magnetized by an inducing field, such as the earths). Topsoil generally tends to be more susceptible due to several processes that include physical and chemical weathering which concentrates magnetic compounds, a fermentation process that changes them to more magnetic forms, and the presence of bacteria that accumulate magnetic minerals (Gaffney and Gater 2003). Although topsoil may be only mildly susceptible, strong anomalies can be formed from it through a number of cultural processes. Abandoned subterranean storage pits and ditches eventually become filled with nearby topsoil by erosion, or were purposefully sealed by the occupants, forming large magnetic contrasts owing to the relatively greater volume of magnetically susceptible material they hold. Mounds built from topsoil likewise produce magnetic anomalies and this is especially true of mounded middens that are composed of highly susceptible materials, like broken ceramics, fire cracked rock, soil from hearth cleanings, and organic matter that promotes growth of bacterial forms that concentrate magnetite (Fassbinder et al. 1990). Stone or sediments imported for architectural constructions can exhibit high (e.g., igneous rocks) or low (e.g., many limestones) magnetic susceptibilities, facilitating their detection through corresponding positive or negative anomalies, respectively. Iron or steel artifacts generate extremely large anomalies, which can be a blessing if one is seeking a few iron-bearing artifacts in a proto-historic site, but a nightmare on certain historic period sites (or sites with modern surface trash) where such items as a rain of nails, steel cans, or wire can obscure subtle magnetic details beneath. Strongly magnetic artifacts or burned elements tend to yield a *dipole* field, expressed as paired positive and negative extremes (much like the north and south poles of a magnet), frequently aligned on a north-south axis unless the source has a long axis aligned in a different direction (see Clark 1990; Scollar at al., 1990; Kvamme 2001; Weymouth, 1986, for other details).

4.1.5. Electrical Resistivity

Resistance of the soil to an electrical current depends on a number of factors including moisture, dissolved ion content, and the structure of soil particles and its components. A resistance survey is an *active* prospecting method that utilizes two probes to establish a current through conductive earth, which is measured. Two other probes measure voltage, and the ratio of voltage to current yields resistance, according to Ohm's Law (Clark 2000). In the traditional Wenner configuration, the four probes are each separated along a line by an equal distance, with the current probes on either end. In a uniform deposit (e.g., an alluvial fan of sand), voltage varies with distance from the current probes in regular hemispheres. If voltage is measured on the surface one meter from a current probe, the value recorded is equivalent to the voltage a meter below the surface, allowing a means to control prospecting depth. In practice, the subsurface does not usually present a uniform matrix. Moist or conductive deposits (e.g., clays) provide

an easy pathway for the current, but it must flow around resistant deposits and objects like rock, which alter the voltage and therefore resistance, creating anomalies. A consequence is that the probe separation or depth criterion becomes only an approximation in complex deposits. A further complication is that the recorded resistance in ohms is dependent on inter-probe distances and configuration. Consequently, measurements are normally converted to *apparent resistivity*, a measure reflecting a bulk property of the ground, not dependent on a particular probe arrangement. For the Wenner array, resistivity in ohms per meter is given by:
$ohm\text{-}m = 2\pi Rd$, where R is resistance in ohms and d is the inter-probe distance in meters (Clark 2000).

Many other probe configurations exist that offer advantages and disadvantages according to target depths of prospecting, resolution of features, and speed of survey. In this SERDP project a *twin-probe array* was uniformly employed for all surveys. This probe configuration is essentially a Wenner array split in half, with one current and one voltage probe held in a rigid frame a fixed distance, d, apart. The remaining current and voltage probes are removed to a remote locus and connected by a cable (a minimum of $30d$ away from the survey to eliminate a proximity effect). Apparent resistivity can be estimated through the following equation (Geoscan Research 1996): $\rho = 2\pi R/(1/d_1+1/d_2)$, where R is the measured apparent resistance in ohms, d_1 is the inter-probe distance in the mobile frame, and d_2 is the inter-probe distance between the remote probes (it is assumed that the remote probes are removed from the mobile probes by at least $30d_1$). This configuration achieves improved and easier to interpret feature definition and greatly increases survey speed because only the two probes held in the frame, rather than four, are moved per measurement station.

Under very dry conditions it may not be possible to promote current flow (solutions are to wet the ground prior to the survey or to insert the probes to a greater depth in order that the current may find a pathway to more conductive earth below). Conversely, significant contrasts between subsurface features may not occur if the ground is completely saturated. Obviously, ground moisture plays a major role in resistivity surveys and markedly different results can be obtained at the *same* site depending on ground moisture conditions.

4.1.6. Electromagnetic Induction

Electromagnetic (EM) induction instruments utilize low-frequency radio waves to actively transmit electromagnetic energy into the ground. The energy causes eddy currents to be generated in subsurface conductors, which in turn transmit a weak secondary electromagnetic field recorded by a receiver within the instrument. These signals contain three important components. The first is the primary signal sent directly to the receiver by the transmitter, which is made null during instrument setup. The second component is made up of electromagnetic energy 90° out of phase with the transmitted signal, known as the *quadrature phase*. It is related to the electrical conductivity of the soil, the theoretical inverse of resistivity (highly conductive objects exhibit low resistivity and low conductivity objects show high resistivity). Conductivity is measured in millisiemens per meter (mS/m), and the theoretical relationship with resistivity is given by: mS/m = 1,000/(ohm/m). The final component is in-phase with the primary signal and is related to the magnetic susceptibility of the soil. It represents the ratio in strength of the

induced to transmitted signals, generally quantified in "parts per thousand" (ppt). It is stressed that the in-phase component responds only to magnetic susceptibility and not thermo- or other forms of remanent magnetism. Results can therefore be very different from magnetometry surveys. Thus, EM instruments are capable of generating two modes of data reflecting important and generally independent dimensions of subsurface deposits (Kvamme 2001).

Prospecting depth in EM instruments is generally controlled by the separation distance between transmitting and receiving coils or by changing transmitting frequencies, but most instruments are manufactured with fixed settings. Normally used with coils oriented in a *vertical dipole mode* (VDM), the coils may be placed horizontally through a 90° rotation to achieve a horizontal dipole mode (HDM), where prospecting depth is approximately halved. The in-phase or magnetic susceptibility component offers a much more limited prospecting depth (less than a half-meter), because the active signal is attenuated going into the ground *and* on its return to the receiver, causing sensitivity to fall off at a rate of $1/d^6$ (where d is distance to target), in contrast to a magnetometer's $1/d^3$ (Clark 2000).

Unlike resistivity surveys, EM instruments are sensitive to buried metals, ferrous and non-ferrous, that show up as extreme measurements (metals are highly conductive). Their general advantage is great speed—without probes to insert in the ground, data acquisition is very fast, but these instruments typically lack the resistivity meter's ability to easily target specific depths through simple adjustments in probe separation. Moreover, in arid or dry landscapes probe contact resistance of a resistivity meter can be extremely high making it impossible to promote current flow through upper dry layers. An EM instrument, on the other hand, can sometimes penetrate to moist layers below, permitting data acquisition (Kvamme 2001).

4.1.7. Ground-penetrating Radar

Ground-penetrating radar (GPR) send rapid but distinct pulses of microwave energy into the earth along the length of a survey transect. These pulses reflect off such buried elements as stratigraphic contacts, walls, house floors, pit surfaces, rubble, or middens. The velocity of radar energy and its depth of penetration vary greatly with soil type and moisture conditions, depending on their relative dielectric permittivities (RDP; essentially, their ability to hold and transmit an electrical charge). Significant reflections are produced only at contacts offering large dielectric contrasts and only when the geometry permits reflections to be sent back to the receiving antenna at the surface (Conyers 2004). It is the return times or echoes of the pulses that allow estimates of depths to subsurface reflectors, and their magnitudes indicate something of the nature of material changes in the ground. The outcome of a series of reflected waves stacked side-by-side mimics a vertical section or profile along the length of a transect, where the horizontal axis represents the transect length in real space and the vertical axis represents reflection times, a proxy for depth beneath the surface (Figure 4.3). The length of a transect represents a collection of traces, with perhaps 20-40 samples per meter. The waveform of each is quantized into samples, perhaps 512, that measure amplitude. Large amplitudes imply greater dielectric contrasts between subsurface materials. If wave velocity, which is a function of the material the energy is passing through, can be determined, then the vertical axis may be converted to estimates of depth beneath the

surface. Several methods exist for estimating velocity (Conyers 2004; Kvamme 2001), the most general being knowledge of soil type (where characteristic velocity values can be looked up in standard texts). Depth penetration is limited by a user-specified time window within which reflections are captured, or attenuation of the energy, whichever comes first (Figure 4.3). Thus, GPR data in their raw form are ideally suited for gaining information in the vertical plane, including stratigraphic relationships.

Obviously, with closely spaced parallel transects software can correlate and interpolate information about reflector strength between profiles, allowing a three-dimensional "cube" of data to be generated. This permits analyses of data not only along profiles (important for stratigraphic investigations), but also laterally in plan view, enhancing interpretability because many archaeological features are more easily recognized when seen in a plan than they might be in cross-section (Conyers 2004). This tactic is known as "time-slicing," because the vertical axis represents time and a slice of time is taken from each profile to form a two-dimensional horizontal composite. Time-slices are used extensively in this project and are described in greater detail in Section 4.4. Thus, GPR potentially offers large potential for subsurface studies by offering detailed information in three dimensions.

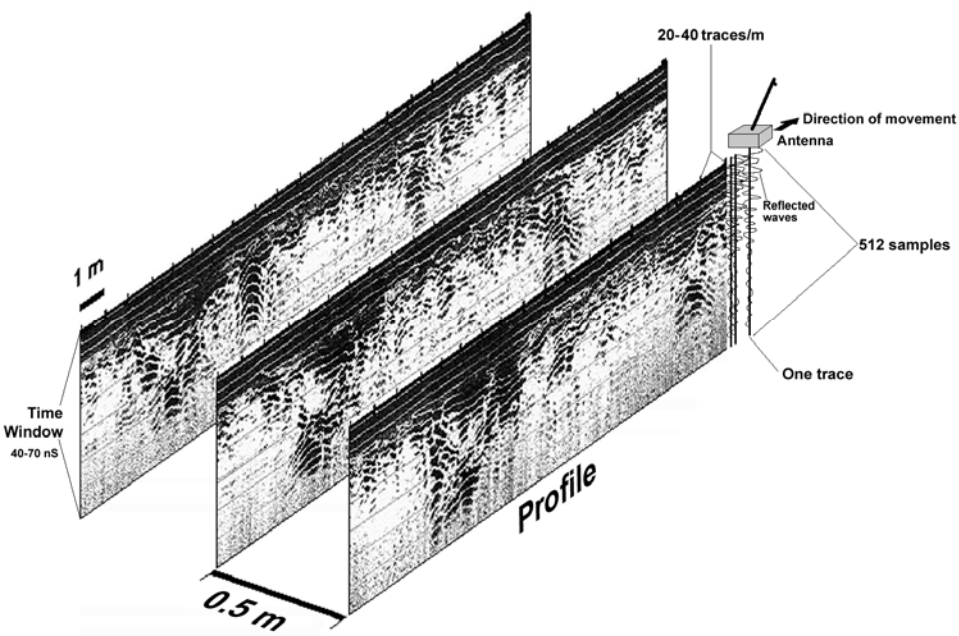


Figure 4.3. *GPR field methodology and sampling, showing profiles composed of traces, digitized into samples within a user-specified time window. GPR reflections are indicated by black wavy lines.*

4.2. GEOPHYSICAL FIELD SURVEY METHODS

Kenneth L. Kvamme, University of Arkansas

With a common archaeological focus on lateral geophysical variations—which enable recognition of culturally formed patterns in anomaly distributions (see Section 2.2)—all data acquisition in this project was gathered through *area* surveys. The point of the area survey is to uniformly sample geophysical measurements over entire regions of study.

4.2.1. Survey Blocks

Ground-based surveys in the SERDP project were conducted piecemeal, generally within 20 m square blocks accurately located by transit or EDM survey, except GPR which utilized blocks of larger and variable size (e.g., 20 x 40 m, 50 x 50 m). GPR blocks were fit within the framework of the 20 m blocks, however. The corners of the 20 m units served as ground control points in subsequent aerial investigations (Section 4.7). Use of small survey blocks allowed work to be broken up into manageable segments, with several blocks targeted each day. Data from the blocks could then be concatenated after fieldwork to yield composites of the full study regions (Figure 4.4a).

4.2.2. Data Acquisition and Sampling Density

Within each survey block surveyor's tapes, with meter and sub-meter marks, were employed to guide the passage of instruments (Figure 4.4b). Instruments are moved along each guide allowing measurements to be accurately located at sub-meter increments. The result is a systematic matrix of measurements with rows paralleling the guides and columns representing the individual measurements in each row. Spatial resolution is controlled by the separation between individual transects paralleling the guides, the number of samples taken per meter, and the sampling capabilities of the technology employed (see Clark 2000:158-164). Electrical resistance meters acquire data slowly (e.g., as slow as 1 measurement in 5 seconds), while EM instruments (2 measurements/second), magnetic gradiometers (8 measurements/second), and GPR (60+ traces/second) are considerably faster.

One project goal was coverage of large areas to facilitate recognition of cultural patterns in the outcome. Another goal was high spatial resolution to allow discovery and definition of cultural features of small size. The former recommends widely spaced samples to insure large-area coverage in limited amounts of time; the latter demands the reverse because sample spacing is the limiting factor that determines the size of archaeological features that can be resolved—data points should be no greater than half the size of the smallest feature to be detected (Section 4.1.3). In this project a compromise survey tactic was sometimes employed, especially for instruments with slower data acquisition speeds. On one axis (usually east-west) moderately high spatial resolution was achieved by employing one-meter transect separations. This required instruments to be walked in only 20 transects per 20 m block. On the orthogonal axis (usually north-south) high sampling densities were employed, usually one-half meter (for slower EM and resistivity surveys) or one-quarter meter (for faster magnetic gradiometery surveys) (Figure 4.4c). This mix allowed very good spatial resolution in one dimension, moderately good resolution in the other, and simultaneously permitted relatively large areas to be surveyed because fewer transects were walked. In some cases half-meter transect separations were employed, allowing much great data densities (Figure 4.4d), especially for surveys by the faster magnetic gradiometer or when sufficient time allowed. All GPR surveys utilized half-meter transect separations owing to the demands of time-slicing methods (see Section 4.4), but 20-40 traces per meter were possible because of the acquisition speed of this instrument.

As results below will show, these relatively high sampling densities allowed many small features to be detected and accurately located in the results, such as narrow

walls, gutters, and hearths. Geophysical field methods are summarized in detail by Clark (2000), Gaffney and Gater (2003) and Kvamme (2001).

4.2.3. Survey Issues

Level fields with short mowed grass make instrument passage easier while large rocks, trees, bushes and other impediments can hinder survey or make it impossible. All of the sites offered favorable conditions for geophysical survey except Pueblo Escondido, which had perhaps 30 percent of its area covered in dense desert brush, bushes, and cactus (and a rattlesnake). Considerable manual effort was required to clear this site of obstructions for survey—perhaps 20 percent of the total effort.

Soil moisture is of critical importance to electrical resistivity and EM conductivity surveys, particularly the former. Too little or too much moisture can lower anomaly contrasts. In very dry surface soils probe contact resistance becomes nearly infinite in resistivity surveys, precluding current flow and the acquisition of data (lacking probes, EM methods can sometimes induce EM fields in moist deposits below the surface). At Army City it was nearly impossible to acquire resistivity data in initial attempts owing to a drought during fieldwork in the summer of 2002; a freak local cloudburst remedied that situation and a complete resistivity survey was hurriedly performed in the course of a few days before the ground was again dry. At Pueblo Escondido the desert climate combined with sandy deposits precluded resistivity survey in the principal study block.

External electromagnetic fields are another issue in geophysical surveys because they can interfere with some instruments, for example, by introducing noise or unwanted artifacts to the data stream. They arise from power lines, electrical storms, radio beacons, cell phones, or pagers. The project experienced only minor problems from cell phone traffic and electrical storms.

The largest environmental difficulty was metallic debris or constructions at three of the project sites. Metal artifacts impact GPR, EM, and magnetometry surveys because they are particularly sensitive to metallic objects (magnetometry only to ferrous metals). While this can sometimes be a benefit in locating artifacts of the historic period (and was instrumental at Army City for locating the town's sewer system), modern metallic litter and constructions are often of such density that survey results can be severely compromised.

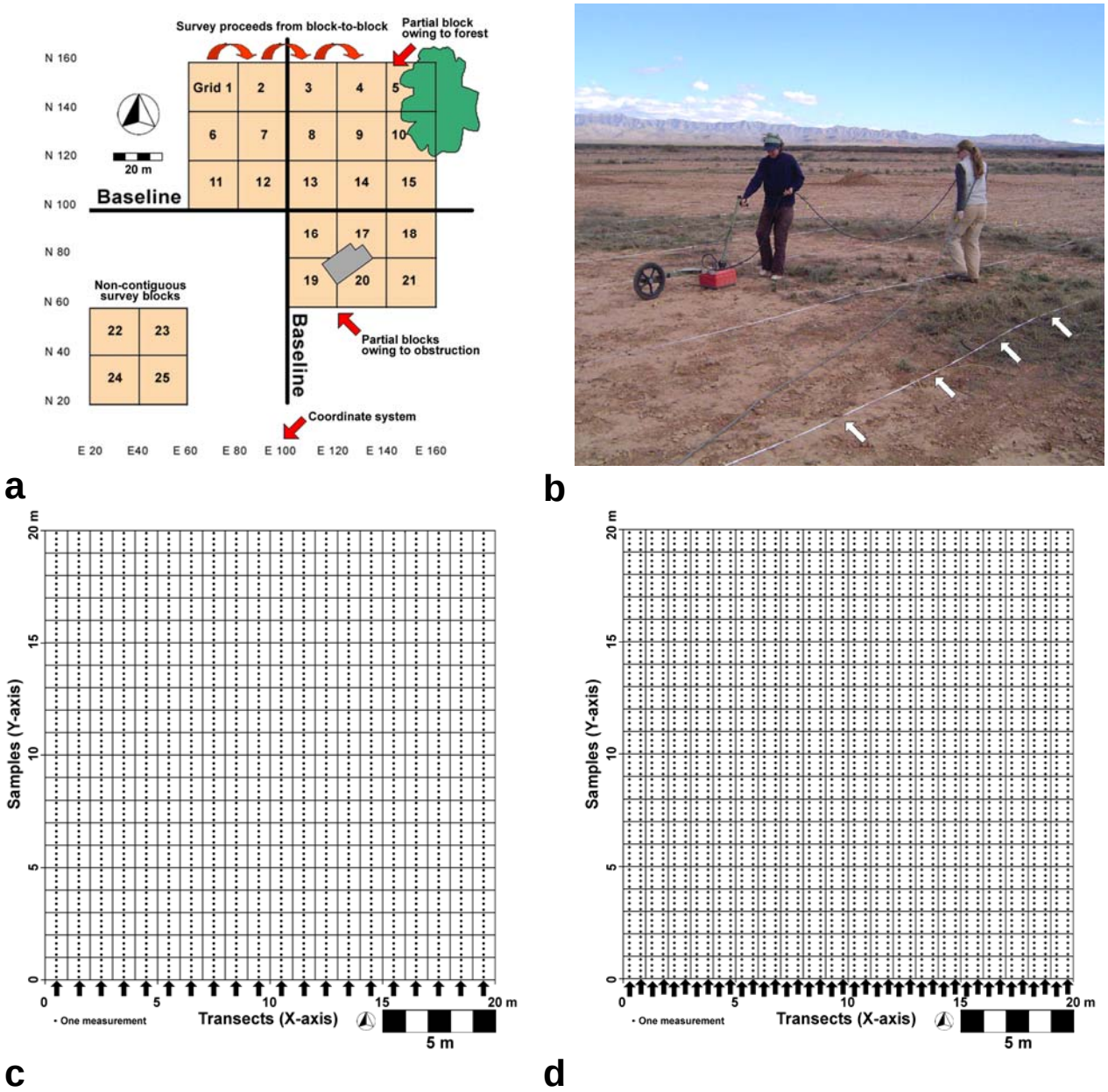
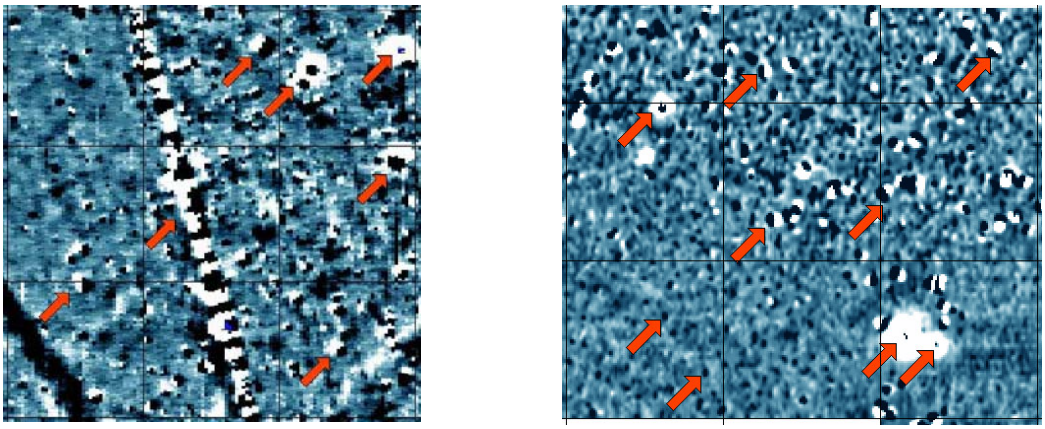


Figure 4.4. *Geophysical survey methods. a) Surveys are conducted piecemeal, usually in 20 m square blocks. b) Survey tapes with meter marks (arrows) are used to guide passage of instrument for accurate positioning in a GPR survey at Pueblo Escondido. c) Sample point distribution in 20 m survey block showing meter-separation between transects (x-axis, indicated by arrows) and quarter-meter samples (y-axis). d) Sample point distribution showing half-meter-separation between transects (x-axis, indicated by arrows) and quarter-meter samples (y-axis).*

Pueblo Escondido had a moderate amount of metallic debris, mostly in the form of shell casings, old fencing wire, and one large rocket imbedded in the earth near the site's center (Fort Bliss is an active area of rocketry). Kasita Village, located adjacent to Lowry Airfield at Fort Benning (active from the early 20th century), was littered with numerous items of metallic debris—apparently composed of nuts, bolts, wire, and other debris from heavy airport use—as well as large buried steel pipelines that crossed the

field for drainage of runways (Figure 4.5a). The biggest problem was at Silver Bluff Plantation. At that site steel wire pin flags had been employed by archaeologists to mark the datum every few meters as guides for future archaeological studies. Unfortunately, prior to our arrival the site was in tall grass and mowed, which resulted in the cutting of hundreds of pin flags and their dispersal across the site (Figure 4.5b). Although brush was too dense at Kasita Village to see metallic debris on the surface, concerted efforts were made to remove such debris prior to our surveys. At Silver Bluff literal bales of pin flag wire were removed prior to our work, but complete recovery was impossible with the result that the magnetic gradiometry data set was compromised (Figure 4.5b).



a b

Figure 4.5. *Effects of ferrous metal debris and constructions on magnetic gradiometry data. a) Pipeline and assortment of dipolar anomalies (arrows) resulting from iron or steel artifacts at Kasita Village, Fort Benning. b) Metal-produced anomalies of large and small size (arrows) arising primarily from cut metal pin flags at Silver Bluff plantation. The grid indicates 20 m survey blocks.*

4.3. GEOPHYSICAL INSTRUMENTATION

Kenneth L. Kvamme, University of Arkansas

Two groups undertook geophysical surveys in this project. The Archeo-Imaging Lab of the University of Arkansas conducted surveys at Army City (Fort Riley, Kansas) and Pueblo Escondido (Fort Bliss, New Mexico). Archaeo-Physics LLC of Minneapolis worked at Kasita Village (Fort Benning, Georgia) and at Silver Bluff Plantation (Savannah River DoE site). Both groups employed identical or very similar instrumentation. This section summarizes equipment used by the project.

4.3.1. Magnetometry

Magnetic gradiometers were employed for magnetometry surveys in the project. They were uniformly carried out with the Geoscan Research FM-36 fluxgate gradiometer (Figure 4.6). These instruments are designed for the rapid measurement of magnetic information over broad areas. As gradiometers they do not measure the total magnetic field strength of the earth below it; rather, they record *differences* between measurements made by top and bottom sensors vertically separated by 0.5 m within the instrument. The bottom sensor is more sensitive to magnetic changes in near-surface soils than the top

sensor (because magnetic field strength falls of with the cube of distance), so the latter primarily records temporal changes in the earth's field. Differencing the two measurements removes these background variations.

Fluxgate technology is very fast and capable of about 0.1 nT (nanotesla = 10^{-9} Tesla) resolution. In magnetic surveys the great majority of the response is from the uppermost meter; more deeply buried features must be highly magnetic to be detected. The instrument may be placed in an automatic recording mode that acquires up to 8 measurements per time unit. The time interval may be varied between about 1.5-3 seconds. Much like a metronome, an audible signal is given at the end of each interval. During survey, the instrument is moved at a uniform pace along a transect with great care to insure alignment with meter marks on nearby survey tapes as each signal sounds (see Section 4.2.2), the speed of which may be regulated according to the operator's walking pace. The matching of the audible signal with meter marks ensures that the data are properly located spatially. The FM-36 is fully computerized and capable of storing 16,000 measurements. This small capacity is limiting, which allowed only four measurements/m to be acquired in transects separated by one meter (i.e., four measurements/m^2).

Figure 4.6. *Magnetic gradiometry survey at Pueblo Escondido using the Geoscan Research FM-36 fluxgate gradiometer.*

Although fluxgate gradiometers are capable of recording data very rapidly, the quality of information acquired is partially a function of how well the instrument is currently "tuned" and how steady the instrument's heading is maintained during a transect—any wobbling or wiggling decreases the signal-to-noise ratio. Fluxgate gradiometers must therefore be tuned or zeroed periodically to reduce this effect and instrument drift. In this process the internal sensors are aligned relative to each other and the earth's magnetic field in *x*- *y*- and *z*-axes such that the directional sensitivity of the instrument is minimized. In general, this tuning process was undertaken at least once for *each* 20 m survey block.

4.3.2. Electrical Resistivity

The RM-15, by Geoscan Research, is specifically designed for the rapid measurement of ground resistance to the flow of an electrical current in archaeological

sites. It is fully computerized and capable of storing 30,000 measurements for later downloading and processing. It utilizes a twin-probe array built into a rigid frame known as the PA-5, which contains current and voltage electrodes (large spikes) linked through a cable to two remote electrodes that complete the circuits (Figure 4.7a). Resistance measurements are automatically sensed and recorded by the RM-15 as fast as the frame can be lifted and inserted in the ground at the next recording station. The focal prospecting depth is controlled by the separation between the PA-5's electrodes, which may be varied in the frame between .25-2 m (in practice there is some variation about the target prospecting depth). In this project, electrodes separation was uniformly set at a half-meter for all sites investigated. Because the RM-15 returns measurements in ohms, it is necessary to multiply them by a constant to convert to an estimate of apparent resistivity (see Section 4.1.5).

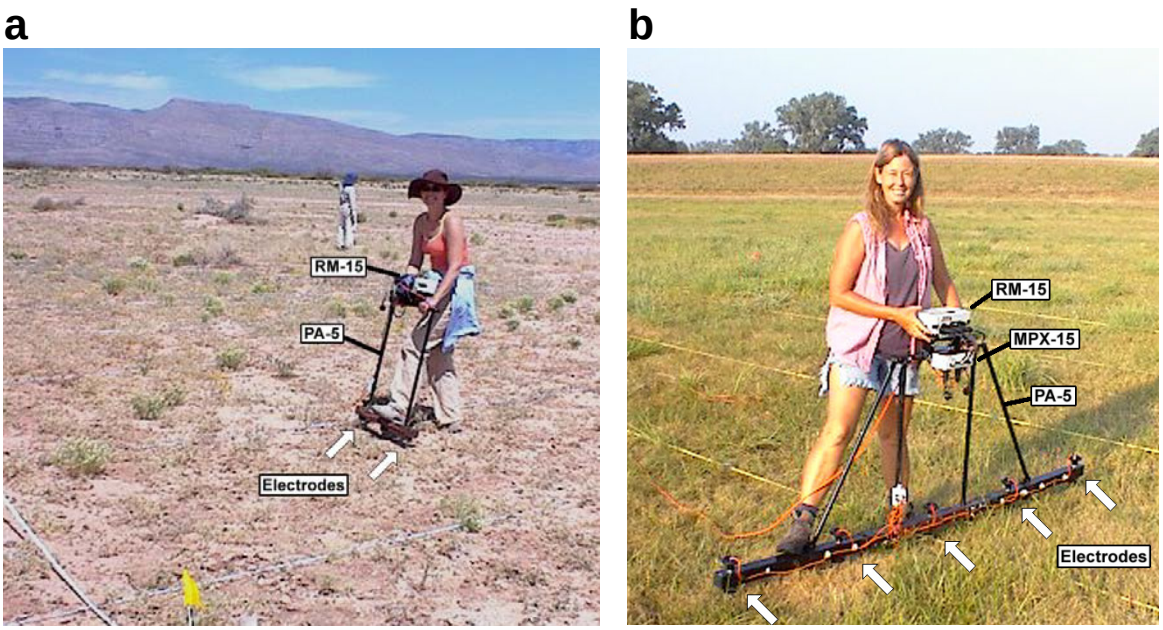


Figure 4.7. *Electrical resistance meters. a) Geoscan Research RM-15 and twin-probe array in PA-5 frame at Pueblo Escondido. b) Geoscan Research RM-15 and twin-probe array at Army City.*

The MPX-15 multiplexer, also by Geoscan Research, is an attachment to the RM-15 that allows near-simultaneous resistance measurements to be acquired from multiple probes at a single station (Figure 4.7b). With the MPX-15 up to six probes at a variety of separation-distances may be attached to the PA-5 frame. The MPX-15 then acts as a high-speed electronic switch that permits recording and logging of multiple resistance measurements between various probe combinations, allowing either (1) multiple-depth data to be recorded, or (2) simultaneous side-by-side measurements at a single prospecting depth. In this project all data were collected with the latter method, which allowed four measurements to be simultaneously acquired with each insertion of the instrument, each separated by a half-meter over a two-meter distance. This allowed very rapid coverage because only 10 transects and 200 insertions were required to completely survey of a 20 m block at half-by-one-meter sampling (two measurements/m^2). Using

only the RM-15 (without the MPX-15), 20 transects and 800 insertions were required per 20 m block for the same sampling density. Although logging of several measurements at a single station with the RM-15/MPX-15 requires more time and the added weight and bulk of the instrument somewhat impedes rapid movement, vastly increased survey speed is achieved. Army City was surveyed by the University of Arkansas' Archeo-Imaging lab with the MPX-15, but surveys at Kasita Village and Silver Bluff Plantation by Archaeo-Physics LLC were performed without the benefit of this instrument add-on.

4.3.3 Electromagnetic Induction

The Geonics Ltd. EM-38 and EM-38B are compact electromagnetic induction meters connected to non-integrated data loggers (Figure 4.8). The Archeo-Imaging Lab used the EM-38 at Army City and the EM-38B at Pueblo Escondido. Both have coil separations of one meter and operate at a frequency of 14.6 kHz. The data logger runs a PC-DOS program that permits complete interaction with the instrument, setting of survey parameters (e.g., sampling intervals, data input rate, acquisition modes), and storage of collected data into files for later downloading to a computer. The EM-38 has two switch-selectable survey modes. The quadrature phase measures soil conductivity as a weighted average through an earth volume of about 1.5 m in a vertical dipole mode (VDM—with coils perpendicular to the ground and peak sensitivity at about 0.4 m depth; see Section 4.1.6). In a horizontal dipole mode (HDM, with the instrument turned on its side) it averages conductivity through a 0.75 m depth (with response decreasing from the surface), but this mode was not used in the project. The in-phase component, sensitive to magnetic susceptibility, offers a much more limited prospecting depth, less than a half-meter in the VDM (Dalan 2006). Because only one mode can be employed at a time with the EM-38, the entire study area of Army City has to be surveyed *twice* to acquire both modes of data, soil conductivity and magnetic susceptibility. This was not the case with the EM-38B at Pueblo Escondido, however, because that instrument allows simultaneous acquisition of both data collection modes.

Figure 4.8. *Electromagnetic induction survey at Army City using the Geonics Ltd. EM-38 electromagnetic induction meter in a vertical dipole mode connected to a non-integrated data logger.*

The EM-38 is capable of recoding data at a rate of two measurements per second, although the newer EM-38B is capable of faster rates. At Army City with the EM-38, data were captured manually by pressing a button on a carrying strap handle (Figure 4.8), but at Pueblo Escondido the EM-38B was used in an automatic recording mode. In that mode a metronome (a low "beep") signals the handler who insures the instrument is placed adjacent to meter marks on nearby survey guide tapes (see Section 4.2.2). The faster EM-38B allowed 4 measurements/m^2 at Pueblo Escondido, while only 2 measurements/m^2 were obtained at Army City.

During survey, the instrument is dragged on the ground to minimize measurement variations stemming from proximity changes to conductive earth. Moreover, it was maintained absolutely vertical in VDM, because any tilting toward the HDM will cause measurement variance. Both of these requirements typically cause about 10 percent of the transects within a 20 m survey unit to be walked two or more times, simply because brush, tall clumps of grass, or other impediments "bump" the instrument during its passage, causing it to lift into the air or tilt dramatically, introducing false anomalies. When this occurs survey of that traverse immediately halts and it is re-walked from the beginning. EM instruments are subject to drift caused primarily by temperature changes. Because of this propensity to drift, the instrument is "tuned" or re-zeroed periodically.

4.3.4. Ground-penetrating Radar

GPR surveys were carried out at all project sites. The Archeo-Imaging Lab utilized a Geophysical Survey Systems Inc. (GSSI) SIR-2000 at Army City and Pueblo Escondido. Archaeo-Physics LLC employed a Sensors and Software pulseEKKO 1000 at Kasita Village and Silver Bluff Plantation. Both have similar capabilities. They are highly portable systems with a control box or computer containing a display screen, antenna, cables, batteries, and other accessories (Figure 4.9). The monitor allows visualization of the subsurface in real time, and the disk drive permits storage of large data volumes. Mid-frequency antennas of 400 MHz (GSSI) and 450 MHz (Sensors & Software) were employed, which offer good spatial resolution and depth penetration to a several meters. Survey wheels were attached to the antennas, which tightly and precisely control the sending of scans or traces into the earth at user-specified sampling distances. The Archeo-Imaging Lab employed 40 traces/m (every 2.5 cm) and Archaeo-Physics LLC acquired 20 traces/m (every 5 cm). A time window or range setting, measured in nanoseconds (1 ns = 10^{-9} second), controls the maximum time allowed for signal return and therefore the depth of prospecting, within a practical limit dependent on soil type, antenna frequency, and other factors (Section 4.1.7). This setting varied from site-to-site, and was usually set to acquire data well below the depth of interest, to 1-2 m. Within each trace, the data are commonly recorded in 16-bit format, allowing over 65,000 levels of amplitude, and each scan is typically quantized in 512 samples. Consequently, 512 x 2 bytes x 40 traces = 40 KB of data might be collected within each linear meter. During data collection, the antenna is dragged across the surface adjacent to a survey guide (Figure 4.9). With a survey wheel, meter marks (Section 4.2.2) may be ignored because the wheel precisely controls sampling. Unlike all other geophysical data sets, data were uniformly collected in transects separated by only a half-meter to allow small targets to

be detected in more than a single pass and improve time-slice visualizations of anomalies (see Section 4.4).

Figure 4.9. *The GSSI SIR-2000 ground-penetrating radar system with control unit, video monitor, keyboard, and battery in the foreground and 400 MHz antenna with survey wheel linked by cable in the background at Pueblo Escondido.*

4.4. GEOPHYSICAL DATA PROCESSING

Kenneth L. Kvamme, University of Arkansas

The computer handling and processing of geophysical data is an essential activity because of the large data volumes generated by modern instruments: a magnetometer might yield 16 measurements/m^2 or 160,000/ha; but even a modest GPR survey will produce a data volume *three orders of magnitude* greater. Computer systems offer the only means to manage and visualize large geophysical data sets. Of equal importance, correct data processing allows subtle patterns to be enhanced and made clear, while defects and forms of noise can be removed or reduced. These factors make the processing of geophysical data nearly as important as collecting the raw data. Although it offers many benefits, it also poses risks. Insufficient processing can leave important anomalies unseen while improper methods might actually remove significant anomalies or introduce artificial ones. This section examines the computer processing of geophysical data. Many methods derive from standard image processing techniques that are common to satellite remote sensing (Jensen 2004; Lillesand and Kiefer 1994; Schowengerdt 1997), but an equal number of procedures are unique to geophysics (Ciminale and Loddo 2001; Conyers 2004; Kvamme 2006a; Scollar et al. 1990:126-204, 488-506). In general, a goal of geophysical data processing is to produce imagery that is as interpretable as possible. In some cases, results can be made to “look like” the buried archaeology such that even a non-specialist can interpret the data, as when the outline of a house is clearly illustrated. This is accomplished by: (1) the removal or reduction of defects and noise and (2) the enhancement of “significant” elements (however significance might be defined). A host

of algorithms or procedures are available for confronting each. Many are reviewed in the following sections.

Any data manipulation must proceed with great care so that anomalies of potential interest are not diminished or eliminated and processing artifacts are not introduced. An excellent precaution is that after each processing step the result is *differenced* from the previous state, allowing examination of changes. These changes reveal what has been removed; if important anomalies are indicated then over-processing might have taken place and that step can be reconsidered.

Procedurally, there is a common series of processing steps applied to geophysical data regardless of source, but GPR data originate in vertical sections or profiles (Section 4.1.7) that require additional pre-processing (described in a separate section). Once GPR profiles are converted into two-dimensional representations of lateral variation they are treatable by the following common methods. Regardless of geophysical data type, information collected from various survey blocks (see Section 4.2.1) are *concatenated* in correct spatial position to make a single composite data set for subsequent processing. Noise and defects are then removed or reduced, followed by enhancements that highlight or exaggerate features of interest. Ultimately, a suitable presentation is designed using various color schemes and forms of computer display. The application of noise reduction or enhancement algorithms is not straightforward, however. Some forms of noise or defects are associated with particular instruments or survey techniques, for example, and some surveys come out relatively "clean" with little need for further processing. The following methods should therefore be viewed as a tool kit to be used as needed to remedy defects, noise, and enhance data subtleties.

A benefit of the on-going computer revolution is that initial data processing may be pursued on-site with portable field computers. This gives an important and nearly immediate link with the data that can influence survey decisions by guiding work to areas of greater archaeological interest and potential. Data processing in this project was accomplished using several software packages, including *Geoplot 3.0* (Geoscan Research), *Radan* (Geophysical Survey Systems, Inc.), *GPR Process* (University of Denver), *Idrisi-Kilimanjaro* (Clark Labs, Clark University), and *Surfer 8* (Golden Software). Much of the following discussion is heavily influenced by capabilities found in *Geoplot* (Geoscan Research 2000) and *Radan* software. General discussions of geophysical processing methods used in archaeology can be found in Conyers (2004), Gaffney and Gater (2003), Kvamme (2001, 2006a), Scollar et al. (1990).

4.4.1. Noise and Defects

Anomalies that represent cultural features are the object of geophysical surveys, referred to as "signal," while "noise" refers to everything else that obscures targets of interest. Frequently, one project's signal may be another's noise, making this process somewhat confusing. For example, iron debris may be a nuisance to the measurement of soil variations in prehistoric sites, but that very litter may represent targets of interest in an historic period site, pointing to significant artifacts. Noise and signal therefore become relative concerns dependent on project goals. Noise derives from three principle sources: elements within the soils and deposits of the site surveyed, from the instrument itself, and from instrument-handling errors by its operator (known as survey defects). The goal of obtaining a high signal-to-noise ratio is facilitated by impeccable field methods, quality

instrumentation that is well "tuned" and correct data processing. The first two can only limit the introduction of survey defects; it is only through data processing that noise inherent to a site can be reduced and signals of interest enhanced.

4.4.1.1. Noise inherent to a site

Noise can arise from such natural disturbances as excessive rodent work, badger or coyote dens, tree throws, and the like. Geological disturbances such as buried paleochannels can produce extensive anomalies as can the archaeological record itself, as when numerous and dense cultural features in higher levels make anomaly detection in deeper levels difficult to visualize. Modern cultural disturbances such as pavements, landscaping, buried pipelines, nearby electrical fields, metallic debris, passing or parked automobiles, trails, or plowing also introduce numerous unwanted anomalies. In general, it is difficult or impossible to remove the effects of large features like paleochannels, pipes, pavements, or tree throws from imagery; they are part of the record. Steps may be taken to eliminate or reduce the effects of small or periodic noise factors, however.

Data spikes refer to isolated extreme measurements. They arise from three sources. Some are caused by physical circumstances inherent to a site. For example, in electrical resistance surveys rocky soil or high ceramic densities may inhibit probe penetration to conductive soils, causing high probe-contact resistance (see Section 4.1.5) and data spikes. In the Army City survey of this project the geophysical surveys were conducted under drought conditions (Section 3.3.4.1), which made the ground highly resistive, causing numerous spikes in the electrical resistivity data (Figure 4.10a). In a similar fashion, metallic objects are highly conductive and ferrous metal artifacts yield large magnetic measurements, causing data spikes in EM and magnetometry surveys. Data spikes may also arise from loose electrical connections within an instrument that introduce uncharacteristically high or low measurements or by operator error. The latter may occur in electrical resistivity surveys when probes are not inserted deep enough in dry soil to promote current flow (Figure 4.10a) or when an EM instrument or magnetometer is moved or tilted dramatically from its normal survey relationship with the ground.

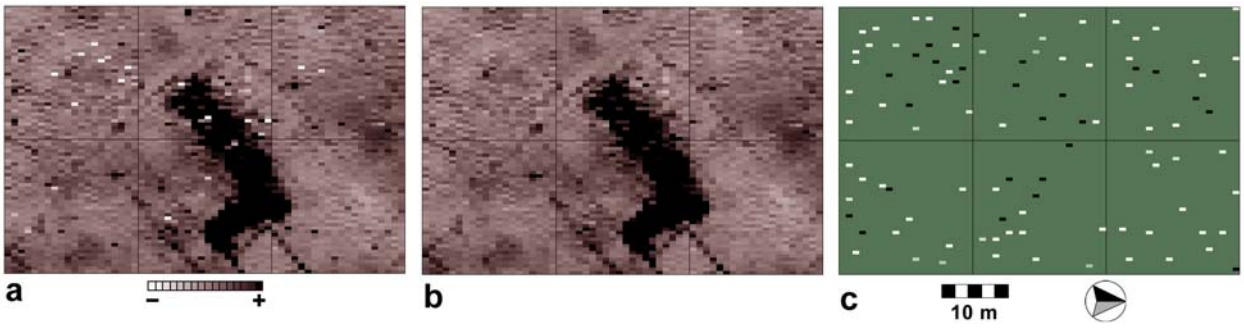


Figure 4.10. *Removal of data spikes, a common form of noise. a) Segment of an electrical resistivity data set from Army City showing numerous data spikes that arose from unusually dry conditions. b) After application of a conservative de-spiking algorithm that removes only the largest spikes. c) Difference between images a and b showing positive and negative data spikes removed.*

De-spiking algorithms generally employ a form of adaptive box filter (Jensen 2004) to remove isolated extreme measurements, typically by replacing them with the *average* of neighboring measurements (Figure 4.10b). This average represents a best estimate of what any measurement should be in the absence of other information. In historic sites, metal artifacts may represent targets of interest. Their removal by de-spiking may improve visualization of soil anomalies, but important information about metal artifact distributions may also be lost.

Plow marks, whether recent or historic, are frequently visible in geophysical data for several reasons. The ridges and furrows represent greater or lesser amounts of magnetically enriched topsoil, causing magnetic variations, or they may be drier and wetter leading to conductivity differences (Figure 4.11a). This type of noise may be removed using Fourier methods (Figure 4.11b,c) owing to their regular orientations and periodic nature (Perkins 1996; Kvamme 2006c). If any linear anomalies of archaeological interest parallel apparent plow marks, Fourier methods will also remove them, so caution must guard against this possibility.

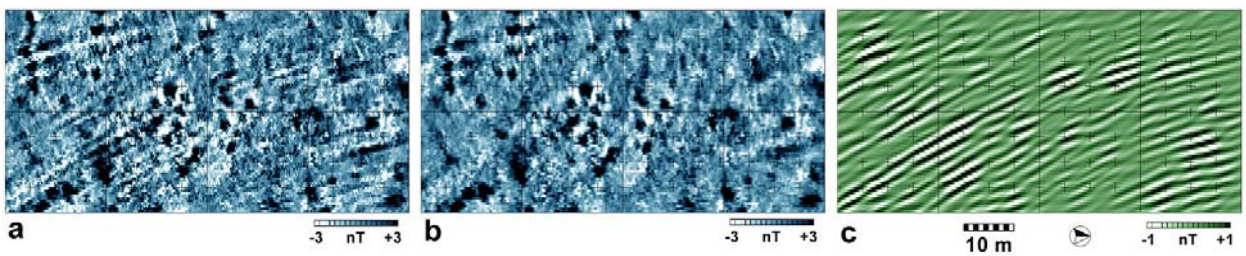


Figure 4.11. *Historic or recent plow marks represent an easily removed form of noise. a) Magnetically revealed plow marks from mule-drawn plowing of the 19th century partially obscure anomalies associated with an early Archaic occupation (ca 7,000-10,000 B.P.) at the Wallace Bottom site (3NW50), Arkansas. b) Fourier methods successfully remove or reduce their presence, allowing improved visualization of other anomalies. c) Difference between images a and b showing image elements removed.*

4.4.1.2. Instrument noise

Instrument noise is inherent to any electronic device and arises from the quality of its electrical components, the nature of the sensors, and how well it might be calibrated or tuned. Such noise is manifested in a number of ways, and a variety of processing techniques can be employed to remove them.

Drift. Some instruments, such as EM induction meters and fluxgate gradiometers "drift," which means their zero-points fluctuate through time owing to perturbations in electronics and other factors such as temperature (Figure 4.12a). Instrument drift causes discontinuities to become apparent between adjacent survey blocks, because adjacent measurements may be of markedly different value. These discontinuities introduce false linear anomalies and must be removed. They certainly detract from image quality (Figure 4.12a). This defect may be handled in various ways. The slope of the trend across any survey block (or sub-block segment) can be determined and removed to "flatten" or remove the trend (Figure 4.12b,c), but sometimes that trend might illustrate real subsurface variations important to the understanding of a site. Commonly, possible drift

in a survey block is examined relative to surrounding blocks. This enables a judgment to be made about which blocks illustrate drift and which represent real geophysical variations (Figure 4.12a). After assessment, de-sloping is applied to appropriate blocks such that measurements along their perimeters closely match measurements in adjacent blocks, yielding a seamless result (Figure 4.12b).

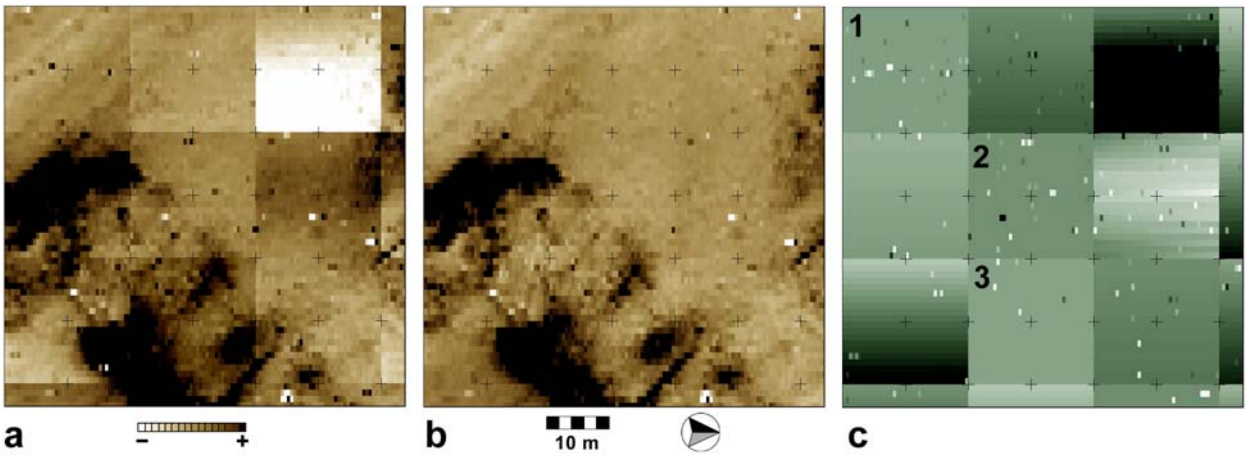


Figure 4.12. *Correction of apparent discontinuities between 20 m square survey blocks. a) EM in-phase (magnetic susceptibility) data from Army City showing block imbalances caused by instrument drift and offset zero points between survey blocks. b) Result of de-sloping algorithm applied to individual survey blocks and block edge-matching algorithm. c) Difference between images a and b showing image changes. Survey blocks 1-3 show no evidence of drift, but indicate unequal instrument zero points. Note that some de-spiking has also been performed.*

<u>*Heading and striping errors*</u>. Drift is also common in surveys by fluxgate magnetic gradiometers, the type of magnetometer used in this project (Section 4.3.1). On top of this, they yield minor "heading errors" that create slight shifts in measured values with changes in the direction it is facing (seen as stripes in Figure 4.13a). This primarily occurs when zigzag surveys are employed, despite operator attempts at constant orientation. Geophysical theory is employed to simultaneously remove drift and heading errors. The theoretical mean of gradiometry data is approximately zero, so these variations can be removed by forcing the mean of each survey transect to equal zero (by adding or subtracting a constant) (Figure 4.13b). In extreme cases, where drift is apparent within *single* transects, a least-squares line is fit to each one. This causes the residuals to be zero-centered (i.e., their mean is zero) which effectively removes drift within or between transects.

Other instruments occasionally yield striped data or heading errors. This is a phenomenon commonly seen in deeper GPR time-slices, caused by processes not well understood. It possibly results from slight but consistent tilting of the antenna such that its "look angle" alternates between the two directions of a zigzag survey (in which case it represents an "operator error"). Striping sometimes occurs in electrical resistance surveys with parallel twin-probe arrays when one of the probe pairs consistently returns higher or lower readings (due to electrical shorts or other factors within the frame; see Section 4.3.2). In both of these cases, zeroing of the transect data as illustrated in Figure 4.13a-c is not theoretically justified and an alternative procedure is warranted.

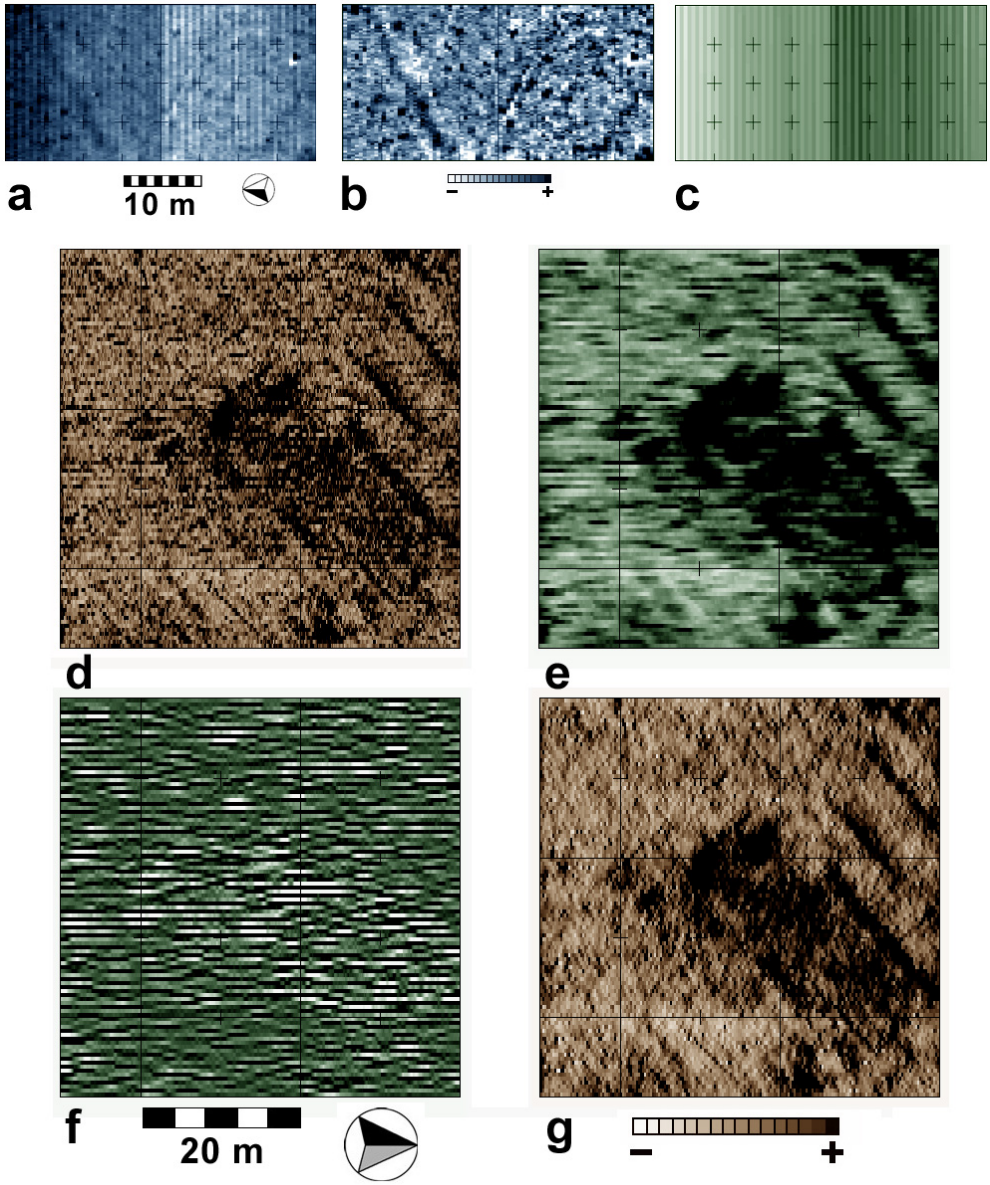


Figure 4.13. *Correction of heading and striping errors. a) Fluxgate magnetic gradiometry data illustrating heading errors and minor drift. b) The gradiometry data after zeroing the mean of each transect. c) Difference between a and b. d) Raw GPR time-slice about 60 cm below surface illustrating moderate striping at Army City. e) Result of large-diameter low-pass filter confined to survey transects (rows). f) A narrow high-pass filter that differences each pixel from adjacent ones above and below; this image characterizes all striping in d. g) The GPR data with striping removed by subtracting d minus f.*

The de-striping algorithm developed by Oimoen (2002) for correcting striping defects common to digital elevation models (DEM) yields excellent results. This method recognizes that striping can be defined as higher or lower average values along a transect (Figure 4.13d). This three-step algorithm first runs a low-pass or averaging filter

(Lillesand and Kiefer 1994) confined to a single transect with kernel length about that of a typical stripe. This step characterizes average values along transects (Figure 4.13e). A 1 x 3 high-pass filter is then applied to the previous result in the *orthogonal* direction. This step computes the average deviance of every pixel in a transect from adjacent pixels in neighboring transects, indicating how much higher or lower a pixel is relative to its neighbors in the low-pass filtered result (Figure 4.13f). The final step subtracts this result from the original to yield a stripe-corrected result (Figure 4.13g).

A danger of any de-striping algorithm occurs when narrow linear archaeological features coincide with the transect direction (e.g., walls, ditches, trails). When the data are normalized, such features can be removed or reduced in the outcome. Care must therefore be taken to ensure that potentially significant anomalies are not removed; if they are, such normalizations should be avoided (or other more specialized steps carried out). Anticipating this possibility, the orientations of all survey blocks in this project were rotated relative to likely directions of lineations at each site (indicated by previous archaeological work or historical sources). This forced survey transects to cross lineations at steep angles permitting de-striping to be safely undertaken (as in Figure 4.13g).

4.4.1.3. Operator errors

A great amount of noise can arise from the instrument operator. The operator must be free of metals when working with EM instruments, and of ferrous metals with magnetometers. Some instruments require a uniform walking speed to insure a constant number of samples per meter and accurate spatial location of the measurements. With magnetometers, a constant instrument height above the ground surface must also be maintained owing to the fall-off of the magnetic field with the third power of distance. EM instruments may also yield variations with instrument height and with how it is angled when carried (see Section 4.3.3). Even with experienced surveyors defects commonly arise from instrument handling, and a variety of data processing algorithms are employed for their correction.

Survey Block "Edge" Matching. Certain operator-caused phenomena introduce survey block imbalances that make edge discontinuities apparent in geophysical surveys. They are particularly common in twin-probe resistivity surveys where remote probes must constantly be moved as a survey proceeds from block-to-block (see Section 4.1.5). Different soil properties beneath these probes at new locations alter apparent resistivity causing block discontinuities (Clark 2000). A simple field procedure exists to correct this imbalance, but minor imbalances nevertheless result, making this a form of operator error. Similar difficulties are also common in EM surveys where slight variations in instrument zeroing cause minor changes in the average value of an entire block, leading to block imbalances. This is illustrated in Blocks 1-3 of Figure 4.12a,c, where instrument drift did not occur but where slight differences in instrument zero points caused somewhat higher or lower mean values. GPR surveys also illustrate block imbalances when rainfall wets the ground in the midst of a project, raising the relative dielectric permittivity of subsequent survey blocks, which alters wave velocities and amplitudes such that discontinuities are apparent between blocks surveyed before and after. Special GPR profile-processing algorithms, discussed below, mitigate this effect, but imbalances nevertheless remain between blocks within time-slices.

A general corrective algorithm that "matches" survey block edges in these contexts determines the mean difference between the edge rows (or columns) of adjacent blocks. This value represents the magnitude of the discontinuity. It is then added or subtracted to *all* measurements in one of the blocks to eradicate the difference and "balance" the measurements such that a "seam" between blocks is no longer apparent (Figure 4.12b). In some cases, for example after a rainfall that alters soil conductivity or dielectric properties, the imbalance not only involves a mean difference, but one of variance. In this case, the *ratio* of the higher to lower standard deviations between adjacent rows (or columns) of two blocks is computed; it is the multiplied against the entire block with lower standard deviation, equating the two. Since the multiplication also changes the mean, it must also be matched by the foregoing procedure.

Staggering. Staggering occurs with instruments that employ timing mechanisms for acquiring data. In this project, this included magnetic gradiometers and EM instruments. With them, the operator must match a metronome-like signal with meter marks on the ground (controlled by tapes placed adjacent to transects, see Section 4.2.2). If timing is off by even a small amount, linear and other anomalies become "staggered," causing a characteristic "herringbone" effect. This is a particular problem in zigzag surveys where the opposing survey directions exacerbate this difficulty (Figure 4.14a). This defect is normally corrected by digitally "sliding" every other transect a small amount up or down (or to the left or right) until the stagger is removed (Figure 4.14b,c). In this project, staggering was minimized by employing experienced operators and primarily running "parallel" surveys with transects walked in a common direction (see Section 4.2.2).

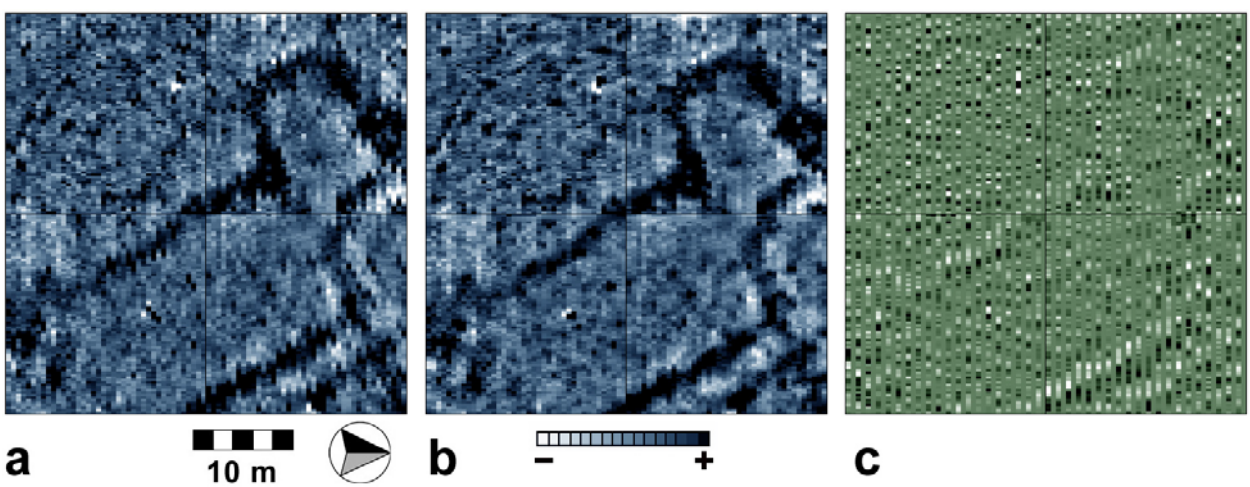


Figure 4.14. *Staggering is caused by operator timing errors, particularly in zigzag surveys. a) Staggering by inexperienced student operators causes a "herringbone" effect along linear features (fortification ditches and bastions) at the Double Ditch State Historic site (ca 1470s-1785), North Dakota. b) The defect is removed by "sliding" every other column (survey transect) a small amount (1-3 pixels). c) Difference of a minus b.*

Gait. With the high instrument sensitivity and sampling rates of magnetic gradiometers, a surveyor's gait can become visible in the data as a periodic defect. This

phenomenon is caused by a combination of (1) regular variations in instrument distance from the surface that correlate with the operator's walking movements, (2) slight angular rotations of the instrument associated with gait (responsible for heading errors, see above), or (3) shoes with mildly magnetic material (e.g., small shoe lace eyelets of ferrous metal or clinging mud of high susceptibility). Because magnetic field strength falls off with the third power of distance, changes of only a few centimeters with each step can introduce a regular sine wave to the data on the order of 0.1-1.0 nT, unless great care is taken to maintain a uniform instrument height. Variations of similar magnitude can occur with subtle instrument rotations, especially as it drifts from a well-tuned zeroing. Magnetic materials on shoes also introduce a sine wave defect with each step. While generally of low magnitude and therefore invisible in data sets of high dynamic range, this defect can become pronounced in magnetically quiet sites with anomalies at the sub-nanotesla level, as occurs at the project site of Pueblo Escondido where the data range of soil anomalies varied between approximately ±1.5 nT. This is illustrated in Figure 4.15a where the operator's gait can be seen as a subtle horizontal banding. Fourier methods are employed for the removal of this defect (Figure 4.15b). The difference between images *b* and *b* emphasizes the periodic nature of this phenomenon (Figure 4.15c). The data also indicate its generally low magnitude, with the full data range of Figure 4.15c varying between only –0.5 to +0.4 nT.

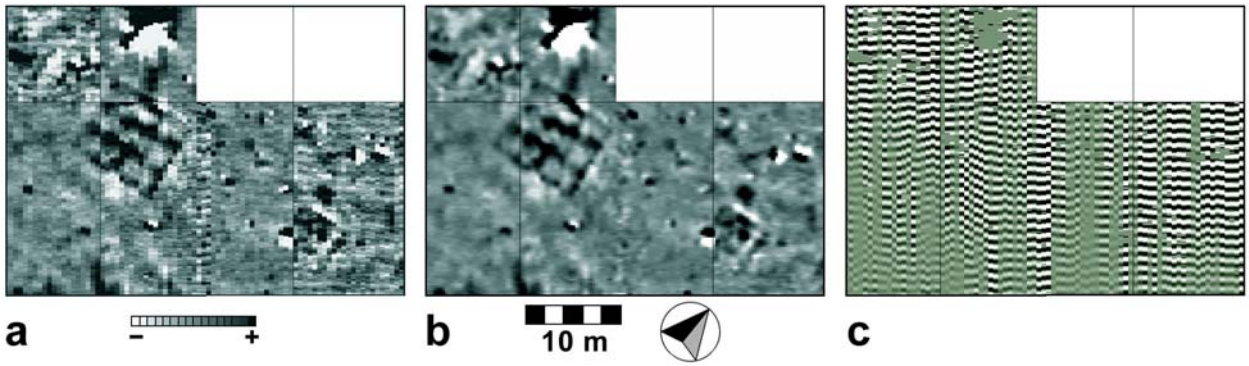


Figure 4.15. *Effects of operator gait on magnetometry data sets. a) Magnetic gradiometry data from Pueblo Escondido showing defects caused by gait; variations in gait magnitude are associate with different operators and how recently the instrument was zeroed. b) After removal by one-dimensional Fourier methods. c) Difference of a minus b showing isolated gait defect.*

4.4.2. Enhancements

Once obvious defects and other forms of noise are removed or reduced from geophysical data sets a variety of enhancements may be applied to better reveal anomalies of interest for analysis purposes or to develop a suitable presentation.

4.4.2.1. Interpolation

Interpolation, the estimation of new measurements from extant ones, is based on the realization that actual field measurements represent a systematic sampling of a *continuous* geophysical field; because it is continuous, estimates of additional values may be made between extant samples using spline or even simple linear functions. Scollar et al. (1990:502) argue that interpolation may be carried out by a factor as large as *4x* without introducing spurious information. Interpolation is employed in this context to

improve image continuity by introducing new measurements that effectively reduce pixel size. This reduces the "block-like" or "stair-step" appearance of pixels and improves image continuity, an important visualization property in gray or color scale imagery. With low sampling densities and relatively few pixels, the eye is drawn primarily to their *edges* rather than to patterns of grays or colors that inform us about image content. That is why one typically "steps back" when viewing such imagery, because apparent discontinuities between pixels then become blurred and data patterns consolidated facilitating recognition of image elements. Additionally, interpolation is used to remedy the effects of unequal sampling densities that typically occur *between* versus *along* survey transects with most geophysical instruments. For example, at Army City, magnetic gradiometry was sampled every .5 x .125 m, GPR every .5 x .025 m, and electrical resistivity every 1 x .5 m. The last is illustrated in Figure 4.16a. These data are interpolated to a uniform .5 x .5 m in Figure 4.16b, resulting in a smoother image where anomalies that apparently represent former walls appear more continuous and are more easily recognized. Interpolation or resampling was also required for data integrations (Section 5) where a common data density was necessary.

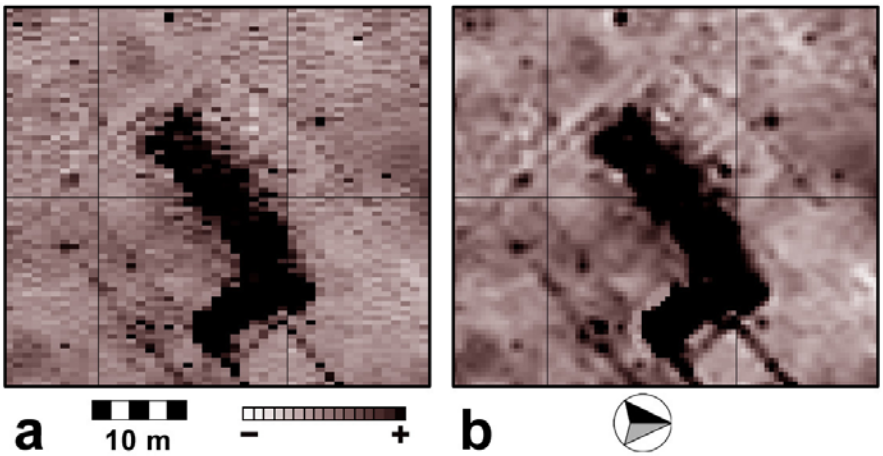


Figure 4.16. *Interpolation improves image continuity, visualization, and the recognition of culturally-generated anomalies. a) Electrical resistivity data from Army City showing anomalies representing an obvious floor and associated walls at the raw field-gathered sampling density of 1 x 0.5 m. b) After interpolation to a uniform .5 x .5 m.*

4.4.2.2. Low-pass filtering

This method is designed to block high frequency information in an image while "passing" only low frequency data, where "frequency" can be thought of as the spatial dimension of image components. For example, random instrument perturbations might occur at a scale of 10-20 cm; rodent holes in a prairie dog town might typically be 30 cm in size; prehistoric ditches are 2 m wide; houses are 6-13 m in diameter; broad geological trends are dozens to hundreds of meters long. Low-pass filtering is designed to reduce the contribution of image elements small in size (here, sub-meter elements), a dimension that corresponds with many sources of noise (Figure 4.17a).

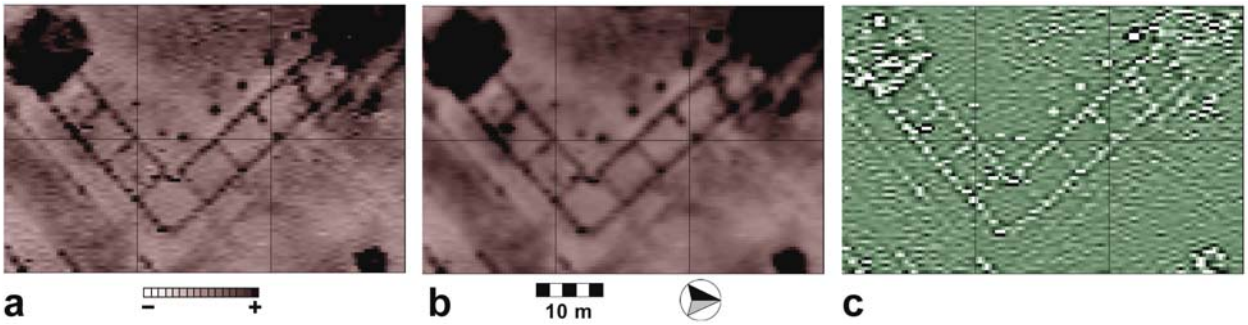


Figure 4.17. *Benefits of low-pass filtering. a) Electrical resistivity data from Army City showing anomalies representing obvious floors and walls after interpolation to a uniform .5 x .5 m. b) The same data after a application of a Gaussian-weighted low-pass filter. c) Subtraction of a minus b showing image changes and lost information.*

At the same time, it tends to consolidate image components of lower frequency, yielding an enhancement of some cultural features by making them appear more continuously and robustly. Low-pass filtering can be accomplished through Fourier methods (where the amplitudes of high frequency components are suppressed or eliminated), but generally convolution techniques are employed (Lillesand and Kiefer 1994). In the latter, a simple average or Gaussian-weighted average of the measurements within a narrow radius or "window" of each pixel is computed (Gaussian weights assign more influence to measurements near the center of the window following the pattern of a Gaussian or normal curve). Simple averaging causes greater smoothing than Gaussian weights as does use of larger-radius windows. A mild amount of smoothing decreases high frequency image components, reduces image noise that might remain after other forms of processing, and consolidates cultural features of lower frequency, making them more apparent (Figure 4.17b). Low-pass filters do cause some loss of high-frequency image information (Figure 4.17c), but this loss is of small magnitude with mild filtering and must be balanced against possible improvements that may result. The data range in Figure 4.17c is less than four percent of Figure 4.17a.

4.4.2.3. Contrast manipulation

Nearly all geophysical data sets illustrate outliers—measurements of extreme value that represent the most robust anomalies. Magnetometry data, in particular, are associated with outliers owing to the presence of ferrous metals (whether modern or archaeological) that generate dipolar anomalies of extreme positive and negative magnitude. These data sets are typically *leptokurtic*, a statistical term meaning that the distribution is highly peaked with long straggly tails (Figure 4.18a). In other words, 95 percent of the data might lie between +/-2 nT, but the remaining five percent can vary between +/- 200 nT or more from relatively few extreme measurements. Obviously, if one "maps" the 256 color or gray values (the common number in most computer displays) to the latter range, only the most prominent anomalies will be illustrated and few of the colors or grays will be assigned to the bulk of the data, resulting in poor image contrast and the invisibility of most features in the data (Figure 4.18a). Contrast improvement is typically achieved through a "linear stretch" with "saturation" of extreme values. That is, the 256 grays or colors are mapped to the range of the central bulk of the data, for example the 95 percent of the measurements lying between +/-2 nT, and the

extreme values are saturated to the lowest and highest grays or colors in the palette (Figure 4.18b). The result is good contrast and image detail because subtle changes represented by only a small range of measurements are portrayed by multiple gray or color values.

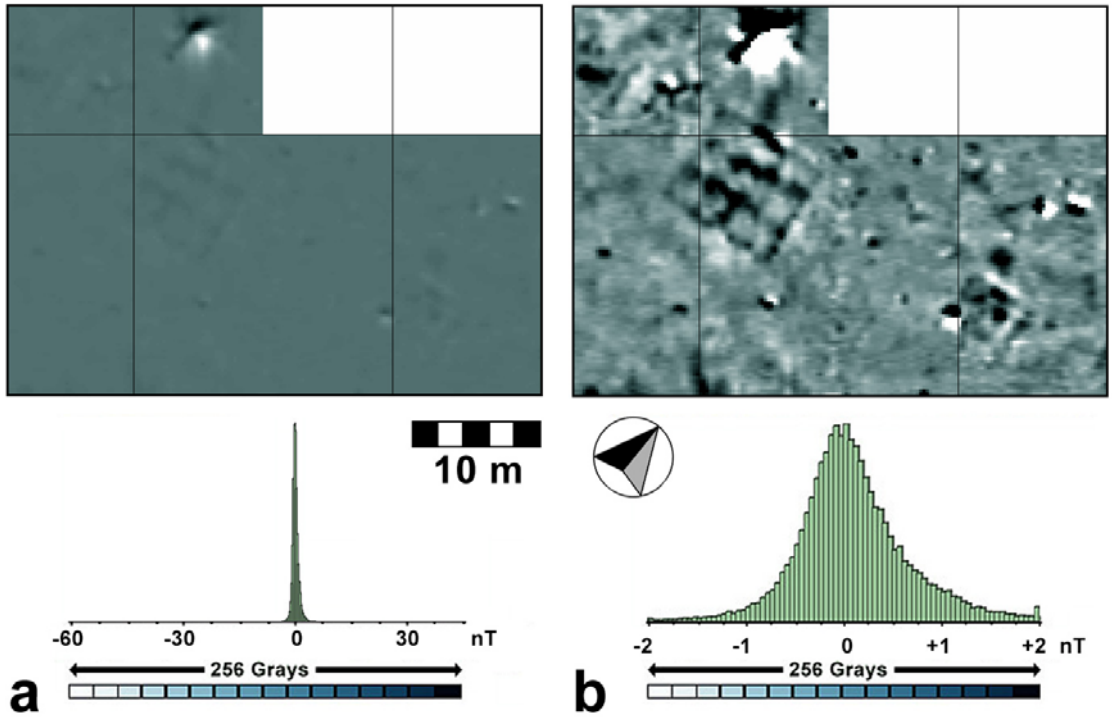


Figure 4.18. *Contrast enhancement of magnetic gradiometry data from Pueblo Escondido. a) Several iron-caused dipolar anomalies (with values ranging from –60 to +40 nT) cause poor image contrast when the 256 available gray values are associated with the full data range. b) A linear stretch that "maps" the 256 gray values to the –2, +2 interval, improving contrast and visualization of image details that reveal pit house structures; data values more extreme become "saturated" with the minimum and maximum display values (white and black, respectively).*

4.4.2.4. Shadowing

Artificial shadowing of geophysical data surfaces can sometimes exaggerate and improve the clarity of subtle anomalies. It is accomplished by computing lit surfaces and shadows resulting from an imaginary light source on a digital surface, such as an array of geophysical measurements. That light source is placed at a low angle above the surface such that "shadows" result from even the smallest variation. Further insights can be achieved by varying the light source direction, because a linear feature is best revealed when the light source is perpendicular rather than parallel to it. Shadowing is illustrated from two directions in Figure 4.19 where a series of "wall" anomalies in electrical resistivity data from Army City are emphasized.

4.4.3. Presentation

The information potential lying within geophysical data may not be realized without a properly designed display that maximizes conveyance of the information it

contains. Modern computer graphics offers a host of ways in which information may be presented. Some more effectively portray geophysical results than others. Choice of display form also depends on the goals of the presentation. Certain forms are best for illustrating relative or absolute differences between the magnitudes of anomalies. Other modes are best suited for illustrating features of extreme subtlety or regular cultural patterns that might occur across an area. Presentation of results to the media or the general public may require yet other kinds of displays.

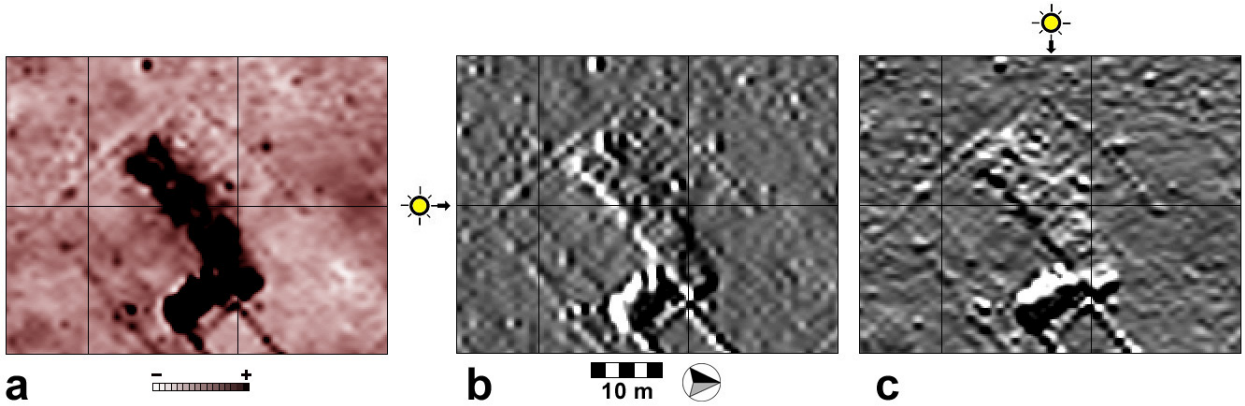


Figure 4.19. *Relief shadowing can enhance image details. a) Electrical resistivity data from Army City showing subtle linear anomalies that probably represent walls surrounding a robust anomaly that likely represents a floor. b) Shadow image with light source from the south and c) shadow image with light source from the west that increase the clarity of certain linear anomalies and yield details within the "floor" area.*

4.4.3.1. Continuous gray or color scales

In general, monochrome imagery best reveals subtle details. In archaeological geophysics, because most anomalies of interest tend to be positive, it is a common convention to illustrate results with what is known as a *reverse* gray scale, with black high (a convention generally followed in this report). With a common eight-bit gray scale of 256 values, extremely subtle anomalies can be illustrated, and systematic patterns across a region more easily discerned and recognized (Figure 4.20). Varied color palettes, on the other hand, can mislead because the eye views transitions between different colors as high contrast changes when, in fact, the actual measurements behind them may vary little. The use of color palettes can therefore obscure subtle anomalies, create false impressions of transitions in the data, and should be employed with caution during data exploration. Yet, a well-designed color scheme can enhance a presentation when anomalies of relatively strong magnitude occur in distinct ranges and specific colors can be assigned to them. Areas that are saturated in gray scale may reveal detail in color (Figure 4.20). Color is also preferred by the media, the lay public, is pleasing to view, and can therefore be essential to project promotion.

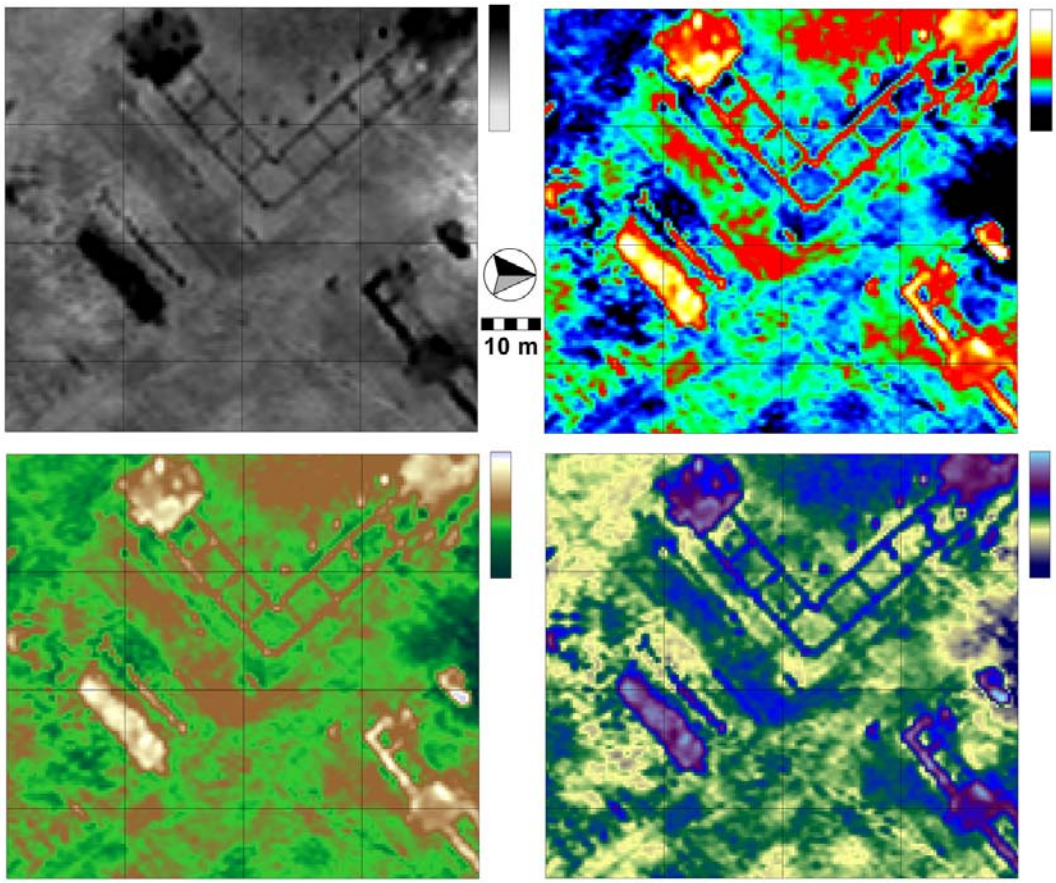


Figure 4.20. *Gray and several color scale displays of resistivity data from Army City where each pixel is assigned a gray or color depending on its measurement value.*

4.4.3.2. Other modes of portrayal

Contouring and pseudo-three-dimensional views of geophysical data offer certain display alternatives to the common gray or color scale (Figure 4.21a). The absolute or relative magnitudes of anomalies can be determined by the sizes or densities of contours (Figure 4.21b) or by how pronounced the "peaks" might be in pseudo-three-dimensional views (Figure 4.21c,d). Such information is typically lost in gray scale because extreme measurements may lie within saturated portions of the scales (e.g., Figure 4.21a). On the other hand, small and subtle anomalies generally become "lost" in contour or wire frame views, and they are frequently ones of high interest. Furthermore, because the eye is drawn to discrete edges, it tends to follow individual contour lines rather than the overall pattern that may be presented. One alternative is to shade a three-dimensional surface representation (Figure 4.21d), which has the advantage of showing subtle detail *and* relative magnitudes of large anomalies at the same time.

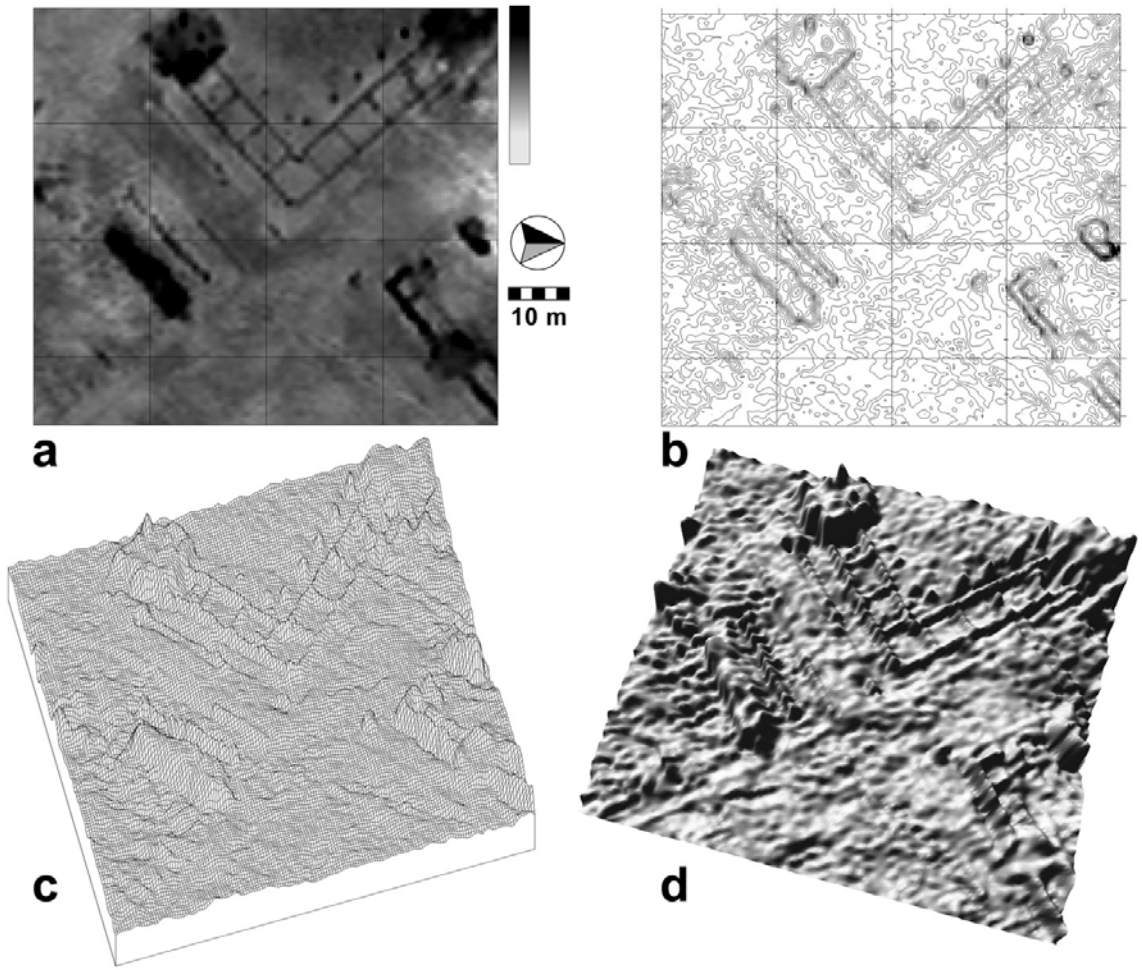


Figure 4.21. *Typical modes of graphic display illustrated with resistivity data from Army City. a) gray or color scale, b) contouring, c) pseudo-three-dimensional wire frame, d) pseudo-three-dimensional surface with shading.*

4.4.4. Processing GPR Data

The foregoing geophysical data processing methods are designed for the manipulation of *lateral* variation from area surveys. In other words, they are most appropriate for two-dimensional data on the horizontal plane. The output of GPR is three-dimensional, however. Each transect of GPR represents a series of reflected waves stacked side-by-side that mimics a vertical section or profile, sometimes referred to as a "radargram" (see Section 4.1.7). With all transects assembled in proper spatial context through software, a true three-dimensional data cube is derived. The horizontal axis of a radargram represents the transect's length in real space, but the vertical axis shows reflection travel time in nanoseconds (see Figure 4.3). The last is a two-way travel time that measures how long it takes each radar pulse sent by the antenna's transmitter to propagate into the ground, reflect off a discontinuity, and return to the antenna's receiver. It is obviously related to depth beneath the surface. The range of the vertical axis is known as the "time window" that determines maximum depth of penetration (although in some cases the available energy may be attenuated by subsurface conditions before reaching the bottom of the window). A series of traces (40 were collected per meter at

Army City and Pueblo Escondido; 20 at the other project sites) lie along the length of each transect that represent reflected waveforms. Each trace is quantized into samples (512 in this project) that give measurements of amplitude or reflection strength. Thus, traces represent columns and samples rows of a radargram.

The processing of GPR data requires (1) treatment of the individual profiles, followed by (2) their assembly in correct spatial position, and (3) the generation of "time-slices"—horizontal plan-view representations of the sub-surface based on lateral relationships between profiles at particular times (depths) below the surface.

4.4.4.1. Processing GPR profiles

Eileen G. Ernenwein, University of Arkansas

Processing radargrams involves a series of steps, including background removal, position correction, gain balancing, and time-to-depth conversion. A great variety of other processes are also used in some cases, but they are too specialized to include here. Due to "system noise," radargrams usually contain unwanted signals in the form of horizontal banding, seen in two adjacent profiles collected at Army City, separated by only 50 cm (Figures 4.22a,b) (Conyers 2004). This banding obscures the reflected waveforms and must be removed. This is accomplished through an algorithm known as "background removal" that improves the visibility of reflections by subtracting the mean of each row (sample) of the radargram from the original data. Figure 4.22c,d shows the previous radargrams after background removal, but the uppermost reflections—representing the ground surface—were not filtered so they could guide the next processing step. The time associated with the ground surface in each radargram should be zero, but the field computer generally offsets it by a few nanoseconds (ns) depending on ground conditions (relative dielectric permittivity [RDP], moisture conditions, coupling, etc.). In order that reflections below will be accurately calibrated against time the surface reflection must be perfectly aligned with "time zero." This adjustment is known as "position correction," and is illustrated for the two profiles in Figure 4.22e,f. The algorithm used in this case also applies a background removal filter, this time applied to the entire radargram to remove the surface reflection.

The two rows of radargrams in Figure 4.22 were collected on different days under *identical* instrument settings, the first under very dry conditions of drought and the second after an unexpected rainstorm. A comparison between Figures 4.22a,c against Figures 4.22b,d illustrates very different contrast, which corresponds with reflection amplitudes. The increased ground moisture has strongly affected reflection magnitudes in the second row of radargrams. The profiles in Figure 4.22e,f were normalized or balanced to each other by a process known as "post-gaining," evidenced by their roughly equal contrast. This processing step is usually only necessary in surveys over which weather conditions change. Gaining is a term that refers to amplification. It was post-applied to all radargrams to modify their signals from top to bottom as a way to balance the profiles, and to compensate for signal loss due to attenuation in the bottom sections of each profile (Figure 4.22e,f). Without this tactic, when the data are examined laterally in plan view through "time-slicing" (see below), radargrams with greater or lesser contrast stand out and must be subjected to an edge-matching algorithm (Section 4.4.1.3).

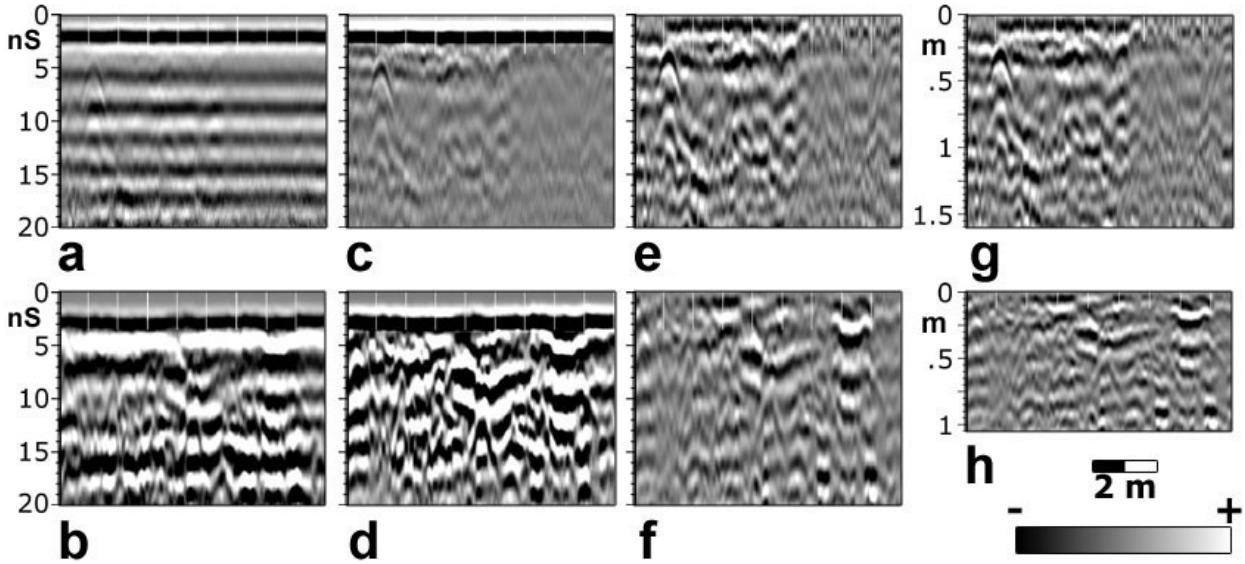


Figure 4.22. *Typical processing sequence for two adjacent radargrams acquired on different days at Army City. a) Unprocessed radargram collected over very dry ground; b) unprocessed radargram collected 50 cm away from "a", when ground was much wetter; c-d) radargrams a-b after background removal of all but the uppermost wavelet; e,f) radargrams c-d after position correction, background removal of the entire image, and gain balancing; g-h) radargrams e-f after vertical scale converted from time to depth.*

4.4.4.2. GPR depth determination
Kenneth L. Kvamme, University of Arkansas

A final step in radargram processing is the conversion of the calibrated reflection travel times to estimates of depth below surface. Radar waves travel at the speed of light in air, but slow down in the ground. Their velocity in the ground depends on its physical properties, primarily grain size and moisture content. Thus, the velocity of radar waves traveling in the ground varies spatially and temporally with changes in weather and ground moisture. Knowing the average velocity for a given location and associated moisture conditions allows the conversion of the vertical scale of radargrams to depth, based on fundamental relationships between velocity, distance, and time. Several methods are available for determining GPR wave velocity. The most accurate is known as the "pipe test" (Conyers 2004). This method requires an open excavation or other vertical exposure into which a metal pipe is inserted at a measured depth (metals are excellent reflectors of radar energy and are therefore easily seen in radargrams). By observing the two-way travel time (TWTT) to the peak reflection of the pipe in the radargram, a close estimate of the average microwave velocity between the antenna at the surface and the pipe may be determined by: $V=2d/s$ (where d=distance from surface to pipe in meters, s=TWTT in nanoseconds, and the constant compensates for the two-way time). Obviously, pipes may be inserted at several depths to investigate changes in velocity with depth, stratigraphy, and moisture content. Depth may be computed from time by: $d=Vs/2$. Unfortunately, no vertical exposures or excavations were available at the SERDP project sites, which precluded use of this method until the final excavation phase of the project (by which time most computer processing was completed).

The more complicated hyperbolic method was therefore employed for velocity determination (Conyers 2004). This procedure arises from two principles, the first being the cone shape of the transmitted microwave signal (perhaps as wide as 60°, depending on RDP). It causes microwaves to reflect from a target *before* the antenna is directly over it. This means that reflection time is minimized only when the antenna is directly over a reflector and that it increases with greater distances from the target. This produces anomalies of hyperbolic form whenever reflector size is small (i.e., "point reflectors"). The second principle is that the width of a hyperbola is a function of the velocity of the microwave radiation, because when energy is traveling very fast it will have a shorter reflection time, creating wide hyperbolas; when it travels slowly reflection times will be longer, and hyperbolas narrower. This property allows velocity estimation because hyperbola shapes can be mathematically described and theoretically calibrated to specific velocities. To illustrate the foregoing, a point reflector (probably a buried metal artifact) was identified in a profile and a hyperbola was fit to it using RADAN GPR processing software. This fit by that program indicates a velocity of approximately .11 m/ns (Figure 4.23). In practice, a series of such fits at about the same TWTT is made and averaged. Additional fits are made to hyperbolas at other TWTT to examine variations in wave velocity with depth.

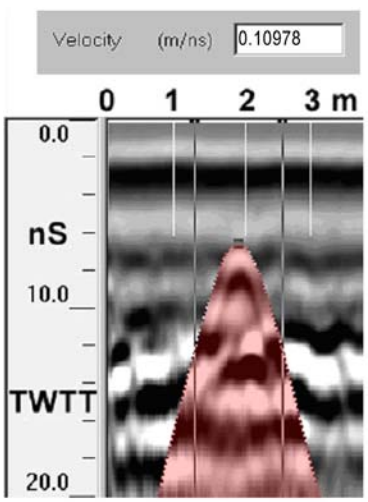


Figure 4.23. *Determining GPR wave velocity. Fitting of a hyperbola to a hyperbolic reflector, which yields a velocity estimate of .11 m/ns based on its width.*

It is well understood that greater soil moisture increases RDP, which decrease the velocity of electromagnetic energy flowing through it by the fundamental relationship $V=C/\sqrt{K}$ (where C=.3 m/ns, the speed of light in a vacuum, and K=RDP) (Conyers 2004). The rain mid-way into the GPR survey at Army City caused a pronounced decrease in GPR wave velocity. The velocity in the profile illustrated in Figure 4.22a, collected under dry conditions before the rain, is approximately .16 m/ns. The velocity in the profile of Figure 4.22b, acquired after the rain, is quite slower, approximately .105 m/ns. These data can be used to recalibrate the vertical scales of the radargrams to estimates of true depth (Figure 4.22g,h), where the unequal vertical lengths of the profiles (1.6 m vs. 1.05 m) within the fixed time windows emphasize the different velocities. The second profile, with its much slower velocity, does not penetrate as deeply as the first. Once this process was applied to all radargrams, "depth-slices" could be created and processed as two-

dimensional images similar to the other geophysical data sets acquired by the SERDP project (as described in Sections 4.4.1-4.4.3).

Incidentally, the estimated velocities allow a determination of the RDP of the soils at the times of the surveys by: $K=(C/V)^2$ (Conyers 2004). The profile in Figure 4.22g yields K=3.5, a very low figure for the silt-clay loam characteristic of the site and reflective of its dry condition in that 2002 summer of drought in which it was collected. The profile in Figure 4.22h yields K=8.2, a much higher value that illustrates the effects of a little moisture.

4.4.4.3. GPR time- and depth-slices

Kenneth L. Kvamme, University of Arkansas

Although GPR profiles are extremely informative, it is difficult to visualize lateral patterns across surveyed regions that allow recognition of regular cultural patterns in anomaly distributions (the "pattern recognition" approach mentioned in Section 2.2.2). In other words, it is much easier to recognize an anomaly as a house in a plan view that illustrates the rectangle of its foundation than in cross-section where complexities of stratigraphy may obscure its form. In recent years, an ultimate goal of GPR data processing has been the generation of horizontal "time-slices" to facilitate the interpretation of GPR data (Conyers 2004; Goodman et al. 1995; Kvamme 2001). With closely spaced parallel transects, such as the half-meter separation distance used in this SERDP project, software can correlate and interpolate information about reflector strength between profiles, allowing a three-dimensional data cube to be generated (Figure 4.24a). This permits creation of horizontal time-slices, which are plan view maps created by taking a particular slice of time from the data cube (with time a proxy for depth; Figure 4.24a). If a *pixel* represents a small area in a two-dimensional image, a *voxel* represents a volume within a three-dimensional data cube. Time-slices, which have an actual "thickness" ranging from a few to many nanoseconds, actually portray a collection of voxels, where it is typically the maximum amplitude or squared amplitude within each that is portrayed. Because GPR data are truly three-dimensional, multiple slices may be extracted, each representing various times below the surface, making it possible to create maps of shallow, mid-range, or deep anomalies (Figure 4.24b). If the vertical axes of radargrams have been calibrated to true depth (e.g., Figure 4.22g,h), then time-slices may appropriately be named "depth-slices," a term sometimes used in the following. Once GPR data are held within depth-slices, they may be processed in the same general manner as other two-dimensional geophysical data sets, as described in Sections 4.4.1-4.4.3.

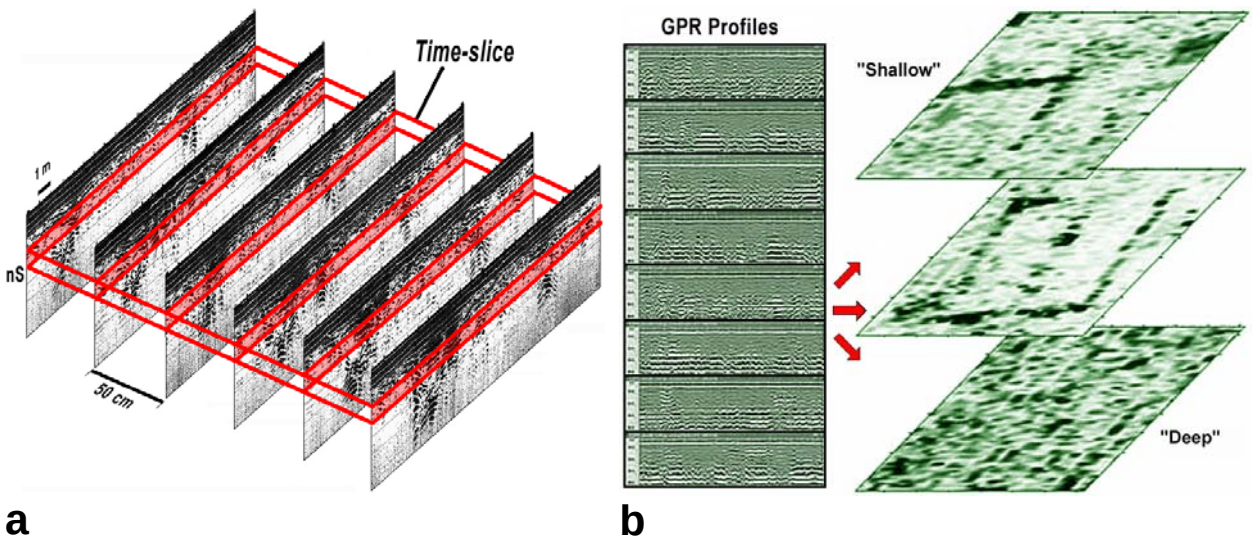


Figure 4.24. *GPR time-slices. a) Illustration of stacked adjacent radar profiles and the "slice of time" (a proxy for depth) to be extracted. b) Plan-view imagery at various times, or depths, below the surface.*

4.4.5. Processing Sequences for Geophysical Data

Kenneth L. Kvamme & Eileen G. Ernenwein, University of Arkansas

Given the wide variety of data processing algorithms for two-dimensional geophysical imagery, great latitude is available in the sequence with which these algorithms may be applied. Yet, it should be clear that some must be applied before others (e.g., de-spiking before interpolation), that some may be unnecessary depending on data type and quality, and that inappropriate application can make imagery *worse* by introducing artifacts and other defects. With the ultimate project goal of integrating or "fusing" the information content of multiple sensors, the data were resampled to common half-by-half meter spatial resolutions from the variety of resolution originally gathered in the field. The following sections summarize data processing sequences employed in this SERDP project using data collected at Army City as an example. That project collected five geophysical data sets, while a fewer number were obtained at all other project sites. Army City can therefore be used to exemplify the nature of geophysical data processing carried out at all sites.

4.4.5.1. GPR processing sequence

GPR processing at Army City first began with pre-treatment of the profile data: background removal was applied, followed by position (time zero) correction, gain balancing, and calibration and conversion from time to depth (Figure 4.22). A series of depth-slices were then generated, each representing likely depths of interest below the surface. Numerous defects are apparent in the raw depth-slice data (Figure 4.25a). Large imbalances remain between survey blocks (despite profile gaining), some striping is apparent, and considerable high frequency noise common to GPR time-slices is clear.

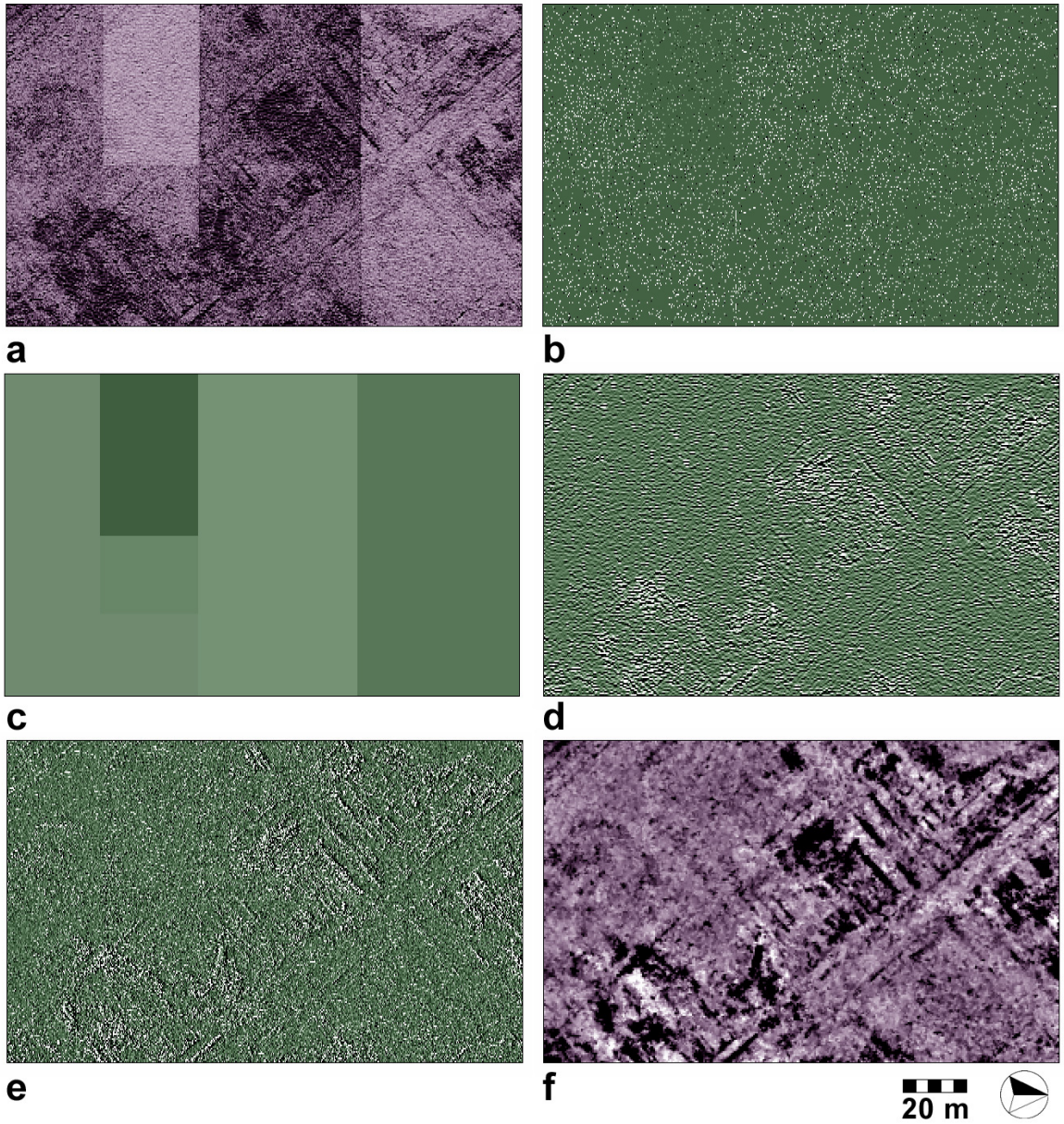


Figure 4.25. *Processing of Army City GPR data. a) Raw depth-slice representing variation between approximately 20-40 cm below surface, which was processed to remove obvious defects, including b) numerous data spikes, c) survey block imbalances, d) striping, and e) additional noise through more de-spiking, resampling to .5 x .5 m, and a low-pass filter, yielding f) the final image.*

De-spiking was first applied because extreme values impact subsequent processing algorithms making it best to remove them first. Edge discontinuities between survey blocks were next tackled, with the goal of producing a seamless image. This was accomplished by matching means and standard deviations between edge rows or columns of adjacent blocks and applying corrective constants to entire blocks, as previously outlined. The lower GPR slices illustrated moderate striping, possibly from antenna tilting (see above). This defect was removed using the low-pass/high-pass averaging

algorithm of Oimoen (2000). To reduce the apparent high-frequency noise that yet remained, further de-spiking and low-pass filtering was applied. Resampling of the depth-slice data from the .125 x .5 m of the depth-slices to .5 x .5 m had a large effect in reducing noise because an averaging algorithm was utilized for this process.

4.4.5.2. Processing electrical resistivity data

The electrical resistivity survey at Army City was conducted under drought conditions and made possible only by a freak thunderstorm that dumped enough rain to enable a quickly conducted survey over the course of three days (prior to the rain, data could not be acquired). The drought conditions nevertheless influenced probe contact resistance, resulting in an unusually high number of data spikes (Figure 4.25.1a).

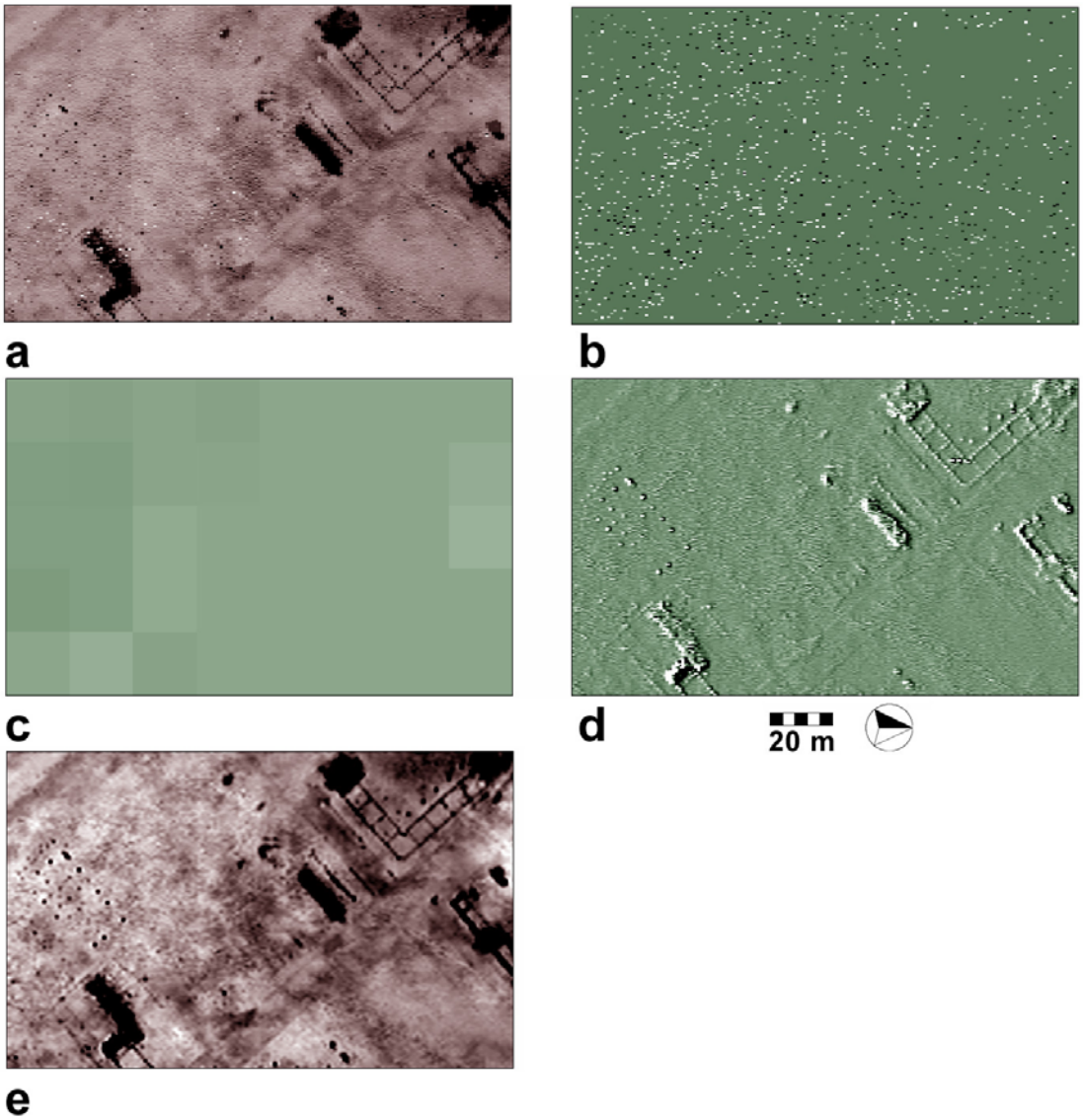


Figure 4.25.1. *Processing of Army City electrical resistivity data. a) Raw data showing several defects, including b) numerous data spikes, and c) subtle survey block imbalances. d) Additional noise was removed through resampling to .5 x .5 m and application of a low-pass filter, yielding e) the final image.*

Remote probes were placed with great care to minimize imbalances between survey blocks in the field; edge discontinuities are therefore nearly absent (Figure 4.25.1a).

Data processing included considerable de-spiking (Figure 4.25.1b), but it had to be undertaken selectively owing to the presence of numerous point anomalies that clearly represented building footings, as suggested by their regular geometric distributions (Figure 4.25.1a). It was imperative that these anomalies not be removed. Minor edge-matching between survey blocks was performed (Figure 4.25.1c), followed by interpolation from the field-collected 1 x .5 m to .5 x .5 m sampling. A mild low-pass filter was applied to consolidate anomalies and reduce image noise (Figure 4.25.1d), yielding the final processed result (Figure 4.25.1e).

4.4.5.3. Processing EM conductivity data

The EM-38 used to collect these data is prone to drift with temperature variations, and temperatures ranged from the upper 60° F in the early morning to over 100° F by mid-afternoon during the surveys at Army City. This caused moderate instrument drift in the quadrature-phase (conductivity) survey, and considerable drift in the in-phase (magnetic susceptibility) survey (see below).

This propensity to drift forced frequent zeroing of the instrument, but minor variations in zero points introduced small survey block imbalances. The nature of Army City and the climate at the time of the survey introduced other data problems. The drought conditions made soil conductivity very low, causing the quadrature-phase survey to respond only weakly to soil variations. Most of the response was to buried metal artifacts, and an early twentieth century site contains a host of them. The conductivity data primarily reveal the buried sewer and water pipeline infrastructure, rebar associated with concrete, and numerous other metal artifacts (Figure 4.26a).

Data processing included de-spiking to remove some of the obscuring noise caused by small metal artifacts (Figure 4.26b), de-sloping of survey blocks for drift correction, and edge-matching between blocks to correct zero-point differences (Figure 4.26c). Interpolation from the field-collected 1 x .5 m to .5 x .5 m was performed as well as the application of a mild low-pass filter to consolidate anomalies and reduce image noise (Figure 4.26d). The final processed image is given in Figure 4.26e.

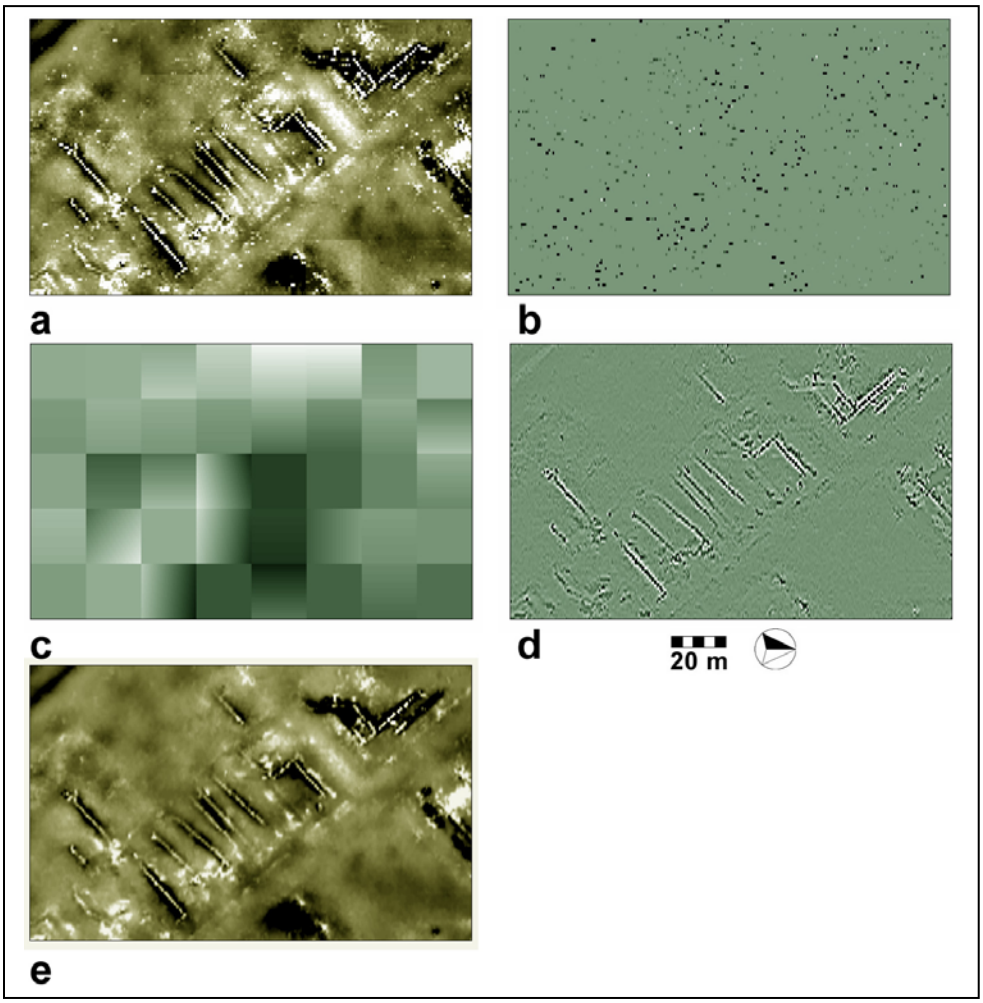


Figure 4.26. *Processing of Army City electromagnetic conductivity data. a) Raw data with obvious defects, including b) numerous data spikes, and c) moderate survey block imbalances. d) Additional noise was removed through resampling to .5 x .5 m and application of a low-pass filter, yielding e) the final image.*

4.4.5.4. Processing EM in-phase (magnetic susceptibility) data

The same instrument used to collect the conductivity data also records magnetic susceptibility (Section 4.3.3), and the same problems tend to occur in this data set, but with minor differences. Drift is considerably greater in the in-phase (magnetic susceptibility) component, but the response of the great number of metal artifacts in the site is much lesser, primarily because of the limited in-phase prospecting depth (less than 50 cm). Buried pipe lines are not visible, for example (Figure 4.27a). The minor variations in zero points caused by frequent instrument zeroing are much more apparent in this phase as survey block imbalances (Figure 4.27a).

Data processing included de-spiking to remove some of the obscuring noise caused by small metal artifacts (Figure 4.27b), considerable de-sloping of survey blocks for drift correction, and virtually every block had to be edge-matched for zero-point differences (Figure 4.27c). Interpolation was applied to change the data density from 1 x .5 m to .5 x

.5 m and a mild low-pass filter was applied to consolidate anomalies and reduce image noise (Figure 4.27d). The final processed data are given in (Figure 4.27e).

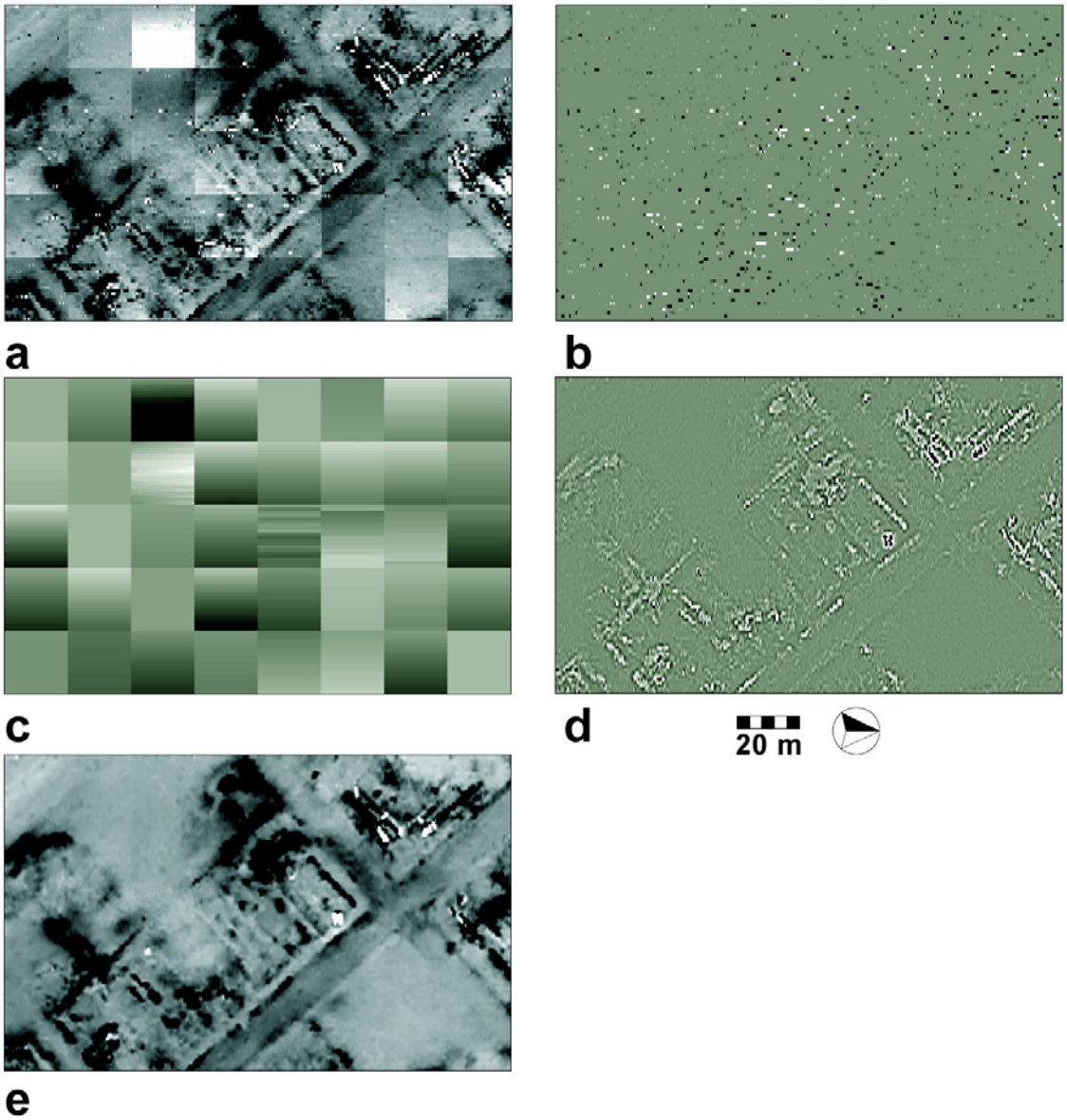


Figure 4.27. *Processing of Army City magnetic susceptibility data. a) Raw data showing pronounced defects, including b) numerous data spikes, and c) large survey block imbalances. d) Additional noise was removed through resampling to .5 x .5 m and application of a low-pass filter, yielding e) the final image.*

4.4.5.5. Processing magnetic gradiometry data

The FM-36 used to collect the magnetic gradiometry data exhibits considerable drift, especially at high temperatures, causing marked slopes in the survey block data and major edge discontinuities (Figure 4.28a). The large amount of ferrous metal artifacts within the site, from pipe lines to rebar and isolated artifacts, introduced a sea of robust and dipolar magnetic anomalies. The magnitudes of these anomalies (hundreds of nT)

generally obscures subtle soil anomalies, but on a positive note, defects arising from gait variations are too low in value to be apparent in the data.

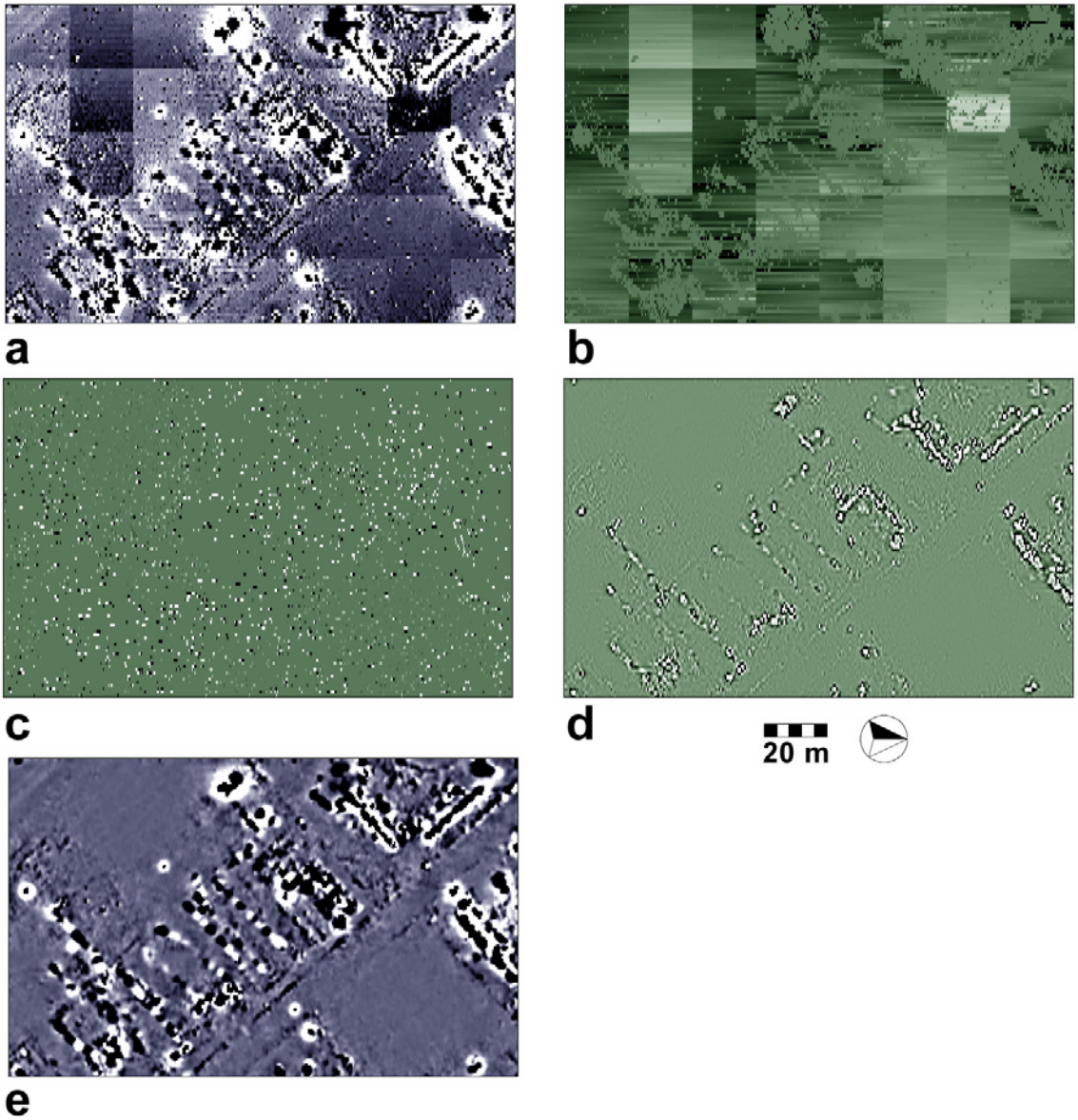


Figure 4.28. *Processing of Army City magnetic gradiometry data. a) Raw data illustrating numerous defects, including b) heavy drift, moderate striping, survey block imbalances, and c) numerous data spikes. d) Resampling to .5 x .5 m and application of a low-pass filter reduces noise and consolidates anomalies, yielding e) the final image.*

Data processing included de-sloping and edge-matching, accomplished through the normalizing of transects by zeroing their means (Figure 4.28b). The last was complicated by the many large-magnitude anomalies, necessitating the isolation of extreme anomalies beyond ±15 nT and the zeroing of sub-transect elements. Moderate de-spiking was also undertaken to remove some of the obscuring noise caused by numerous small metal artifacts (Figure 4.28c). Resampling was next carried out,

converting from the 1 x .25 m field-collected data density to .5 x .5 m, followed by mild low-pass filtering to consolidate anomalies and reduce image noise (Figure 4.28d). The final processed data appear in Figure 4.28e.

4.5. FORMING INTERPRETATIONS FROM REMOTE SENSING DATA

Kenneth L. Kvamme, University of Arkansas

Identification of the source of a remotely sensed anomaly, whether revealed by ground-based geophysical or aerial-space methods, is accomplished through three general approaches. Excavation is a straightforward method, where archaeologists dig at the locus of an anomaly to expose it and identify what it might represent. A second approach utilizes *pattern recognition* (first discussed in Section 2.2.2), where telltale shapes, distributions, and associations give clues about an anomaly's identity. A third technique utilizes the considerable body of available geophysical and remote sensing theory to *deduce* the composition and identity of an anomaly. The following sections overview these approaches to anomaly identification.

4.5.1. Prior Assumptions

Use of any interpretational method for anomaly identification necessarily presumes some knowledge of archaeology in general, and of the local archaeology in particular. In other words, someone *not* trained in archaeology and having no knowledge of what kinds of archaeological features might be expected will be naïve about possible forms in that record and cannot form accurate interpretations. To interpret an anomaly as a storage pit, post hole, or kiln means that one must *know* that these kinds of features are possible and that they might have existed in the culture, area, and time period under examination. Knowledge of archaeological forms, sizes, common associations, distributions, and material compositions is also essential, and obtainable from published accounts of prior excavations in similar sites. At the same time, prior assumptions about what kinds of cultural features might be expected in a site colors interpretations. Being told that a site is historic Euroamerican may force one to zero-in on the rectangular and linear anomalies common to those sites and interpret all such anomalies as Euroamerican—even the prehistoric square house from an earlier occupation. Moreover, non-rectilinear forms may be overlooked that derive from earlier occupations. These notions suggest the importance of experience and caution when approaching anomaly identification.

4.5.2. Anomaly Identification through Excavation

The most direct method of anomaly identification requires excavation and is sometimes referred to as "ground-truthing" (Hargrave 2006). In this method an archaeological team excavates the area in or around an anomaly in an attempt to locate its source and identify what it represents. When exposed in the ground, it frequently is possible to ascertain that an anomaly represents a wall, floor, hearth, pavement, storage pit—or gopher hole, rock, or tree root. Although generally a fruitful method, in some cases the source of an anomaly is no clearer in excavation than it is in remote imagery (e.g., architectural elements or an observable change in soil type, color, or texture may

not be seen). This issue realizes intense focus in the validation portion of this report (Section 5.13).

Excavation as a means of anomaly identification is not generally pragmatic because it is slow, expensive, and a complex archaeological site or settlement may indicate hundreds or even thousands of anomalies. Moreover, excavating each one to identify its source defeats the purpose of remote sensing—to say something about the buried archaeology without having to dig it up! More commonly, a *sample* of anomalies of a particular type and form are excavated to identify what *they* represent; it can then be argued that other anomalies of the same type and form represent similar archaeological features. This technique forms a type of pattern recognition, however, discussed in detail in the next section.

4.5.3. Anomaly Identification through Pattern Recognition

Remote sensing specialists utilize several tactics to distinguish between culturally generated and naturally produced features in the landscape. A primary one is the realization of their *pattern*, which includes characteristics of form, size, distribution, and context. These "recognition elements" were originally codified in the early twentieth century as aids to air photo interpreters (Avery and Berlin 1992), but their applicability extends beyond that limited domain to any field of remote sensing, as the following sections demonstrate (see Kvamme 2006b for a more thorough discussion). As argued in Section 2.2.2, working in this domain demands large-area surveys in order that complete archaeological features and their associations can be seen.

4.5.3.1. Pattern recognition based on shape

A central principal of air photo interpretation is that regular geometric shapes tend to indicate the work of humanity (Avery and Berlin 1992:52). These forms—circles, squares, rectangles, straight lines—occur much less frequently in nature, although natural phenomena sometimes possess distinctive shapes (e.g., alluvial fans, floodplain meanders, volcanic cones). On top of this, cultural objects such as walls, houses, plazas, and roads, usually posses distinct boundaries, unlike natural entities that tend toward irregular shapes and "fuzzy" edges. Consequently, if one sees a rectangle in remotely sensed imagery it most likely represents a human construction. Depending on scale, it could represent the floor plan or foundation of a house, a larger administrative building or monument, a plaza or courtyard, or even the boundaries of an agricultural field. This principle is probably the most fundamental one in the domain of pattern recognition. It frequently allows classification of anomalies as particular types of cultural features without the need for excavation.

Examples of the recognition of archaeological features based on shape are easy to come by. The Great Pyramids at Giza are instantly identified from space, even at the very coarse spatial resolutions available in the 1970s (Quann and Bevan 1977), but recent one-meter imagery from the Ikonos satellite also shows the many rectangular tombs (*mastaba*) surrounding them (Figure 4.29a). Likewise, the locus of the historic Mount Comfort Church (1840s-1863), burned during the American Civil War in Arkansas, is revealed by its rectangular shape (Figure 4.29b), as is the series of shops within the complex know as the Hippodrome, at Army City, Kansas (Figure 4.29c), in electrical resistivity data sets.

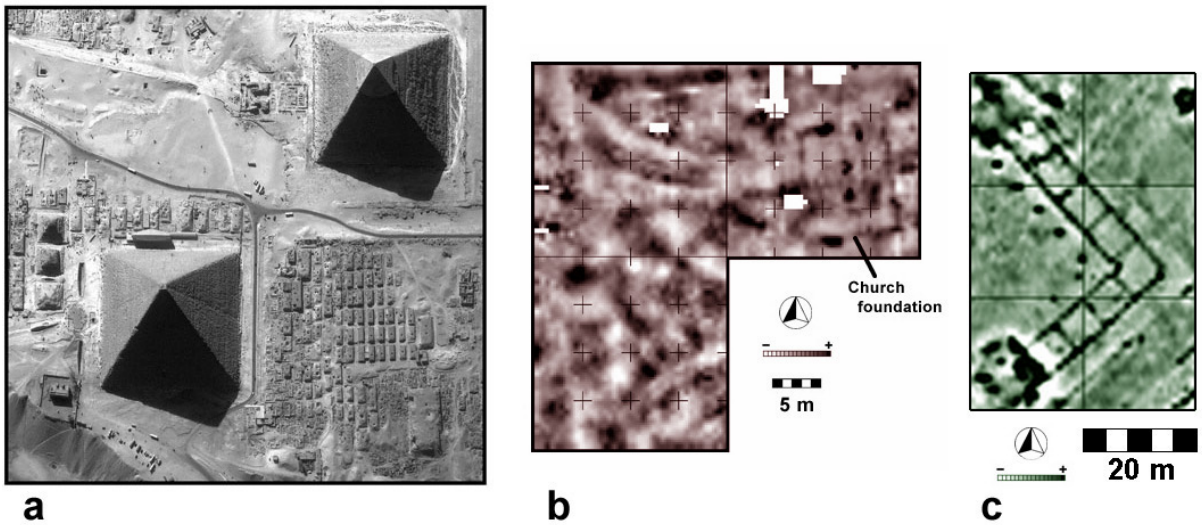


Figure 4.29. *Cultural features exhibit regular geometric shapes. a) The Great Pyramids at Giza, Egypt, are recognized by their triangular faces in this 1 m resolution Ikonos image, as are nearby rectangular tombs (credit: Space Imaging; collected November 17, 1999). b) The rectangular footprint of the foundation of Mount Comfort Church (1840s-1863), destroyed during the Civil War in Arkansas, is shown in an electrical resistivity data set. c) Rectangular rooms of the Hippodrome complex at Army City, Kansas (1917-1921), revealed by electrical resistivity.*

4.5.3.2. Pattern recognition by systematic repetition

Cultural patterns are expressed in other ways. Seen from the air, trees in forests are irregularly distributed, but they occur in systematic rows and columns in orchards. Similarly, in remotely sensed imagery, a collection of small anomalies 1-2 m in size can represent anything from badger dens to tree throws or prehistoric storage pits. When occurring in regularly spaced rows and columns in an electrical resistivity data set, one might make a strong argument for graves, as in the mid-19th century Bozeman cemetery, near Arkadelphia, Arkansas, where few burials are actually marked on the surface (Figure 4.30a). Likewise, when small resistivity anomalies define the perimeter of a rectangular space in an early twentieth century village, the presence of building footings is suggested (Figure4.30b). The latter were recorded at Army City, Kansas (1917-1921), for which a photograph of the structure that was once supported by these footings exists (see right background, Figure 3.2).

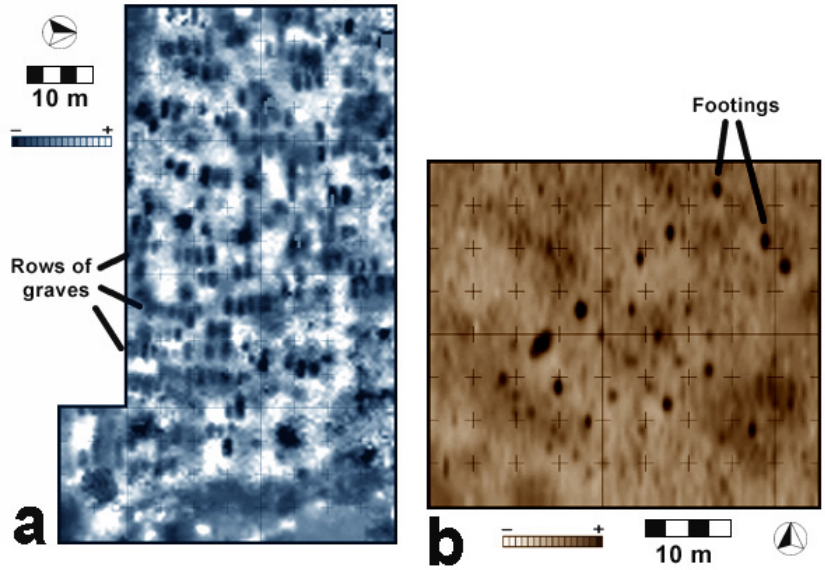


Figure 4.30. *Cultural features may exhibit repetitive patterns. a) Repetitive occurrences of 1 x 2 m oval-to-rectangular shaped anomalies in an electrical resistivity data set point to unmarked graves at the Bozeman Cemetery, Arkansas. b) This regular pattern of electrical resistivity anomalies—forming a rectangle—points to building fooings at the World War I era commercial center of Army City, Kansas.*

4.5.3.3. Pattern recognition by relative size

Size differences are common between many types of archaeological features. Relative or absolute sizes can therefore aid in the identification and interpretation of anomalies. Public buildings versus dwellings, dwellings versus smaller outbuildings and privies, and roads versus trails provide examples of how functionality might be inferred from size characteristics. This phenomenon is illustrated in Figure 4.31a-c where relative sizes between the footprints of a pioneer period cowboy cabin of the 1870s in central Kansas (Kvamme 2001), the previously described Mount Comfort Church (1840s-1863), in Arkansas, and the Fort Clark Trading Post (1832-1861) of the American fur trade in North Dakota (Kvamme 2003) are compared. Obviously, the last cannot represent a dwelling, while such a conclusion is entirely possible given the size of the first. The size of the structure indicated in Figure 4.31b might be too large for a typical house of that period and region, but it is probably normal for a meeting house.

Size differences may be regarded only relatively, but actual dimensions may also give clues relevant to interpretation. A hearth is about a meter wide while most postholes are sub-meter in size; dwellings are at least several meters in diameter. Magnetic gradiometry data from a prehistoric house at Whistling Elk village (ca AD 1300), in South Dakota, reveals a series of interior anomalies that provide illustration (Figure 4.31d). The large central one (about 1.5 m in diameter) was interpreted as a hearth; four small anomalies (about 25-50 cm in size) arranged in a regular quadrilateral were interpreted as likely roof support posts (an inference based on prior archaeological knowledge from similar excavated houses); the bounding square (10 m wide) could only represent the house perimeter. All were made visible by intense firing as the house

burned or through normal hearth use, inferences later validated by excavation (Toom and Kvamme 2002).

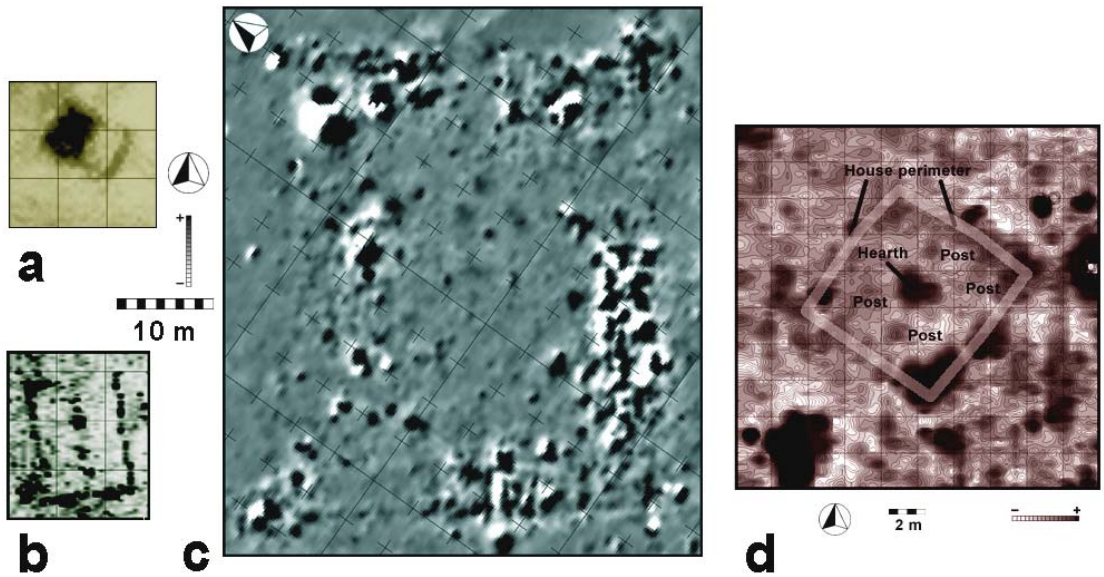


Figure 4.31. *Relative size differences can be associated with functionality, as suggested by comparisons between (a) an 1870s cowboy cabin in central Kansas revealed by electrical resistivity, (b) the mid-19th century Mount Comfort Church, Arkansas, portrayed by ground-penetrating radar, and (c) the Fort Clark Trading Post (1832-61), North Dakota, with magnetic foundation stones revealed by magnetometry. d) A square structure showing relative size differences between its burned perimeter, central hearth, and roof support posts in a magnetic gradiometry image from 14th century Whistling Elk Village, South Dakota.*

4.5.3.4. Pattern recognition by association

Association refers to relationships *between* culturally produced anomalies or features. At a small map scale houses combine to form settlements, fortifications and plazas go with villages, and roads and trails emanate from them, for example. A single house-size anomaly in a field could represent almost anything, from a large tree throw to a silted-in pond, but combine a series of such anomalies in a relatively tight cluster and the probability of them representing houses within a village increases (Figure 4.32.). At Double Ditch village, North Dakota, the association between the fortification systems and the village's dwellings is clear (Figure 4.32b). Associations also occur at a larger map scale. Hearths go with houses as do roof support posts, as occurs in the house at Whistling Elk (Figure 4.31d).

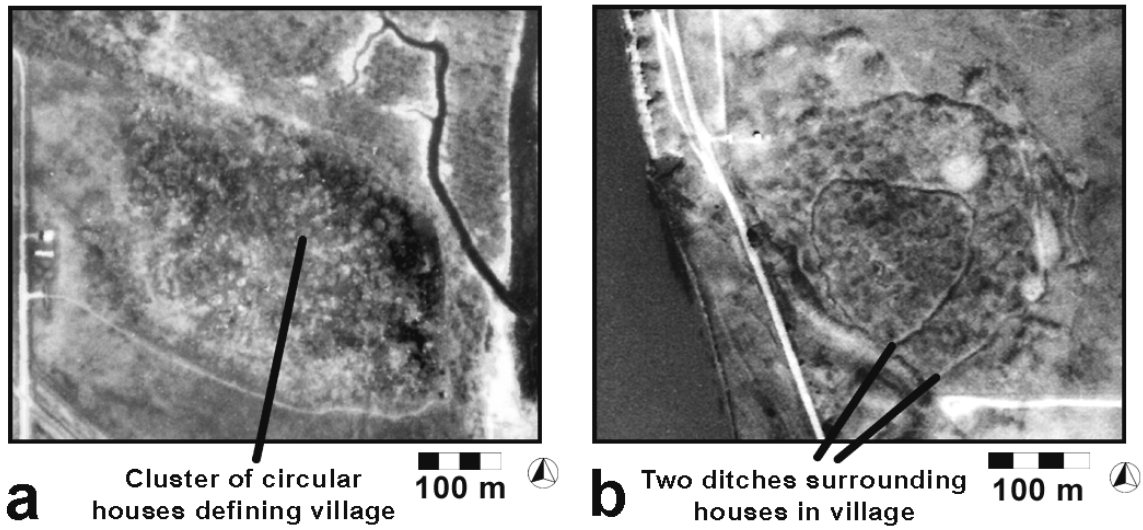


Figure 4.32. *Pattern recognition by association at small map scales in vertical aerial photography. a) Cluster of circular anomalies representing houses form a village (Fort Clark village, North Dakota, 1822-1861). b) Fortification ditches (linear features) are associated with villages (Double Ditch village, North Dakota, ca 1450s-1785).*

4.5.3.5. Pattern recognition by context

The idea of utilizing context as a means to enhance anomaly identification has not been well explored in archaeological remote sensing. Context, as used here, refers to linkages between cultural practices and environment. Frequently, there may be relationships between environment and types of archaeological sites or features. This premise has been well tested archaeologically, with countless locational and spatial analyses performed, and predictive archaeological location models have been developed that capitalize on these relationships (Kvamme 2006c; Wescott and Brandon 2000). For example, villages might be located near the confluences of rivers; defensive sites might be confined to hilltops (e.g., the classic Iron-Age hillforts in much of Europe); and farmsteads to rich agricultural lands.

Context could possibly aid remotely sensed identifications. Techniques could be borrowed from archaeological predictive location modeling where correlations between environmental and site types might help to identify them (Wescott and Brandon 2000). Most settlements of the Plains Village Tradition of the Middle Missouri River tend to be located on high terraces adjacent to the river, for example (Figure 4.32). The within-site use of context may be more problematic. Perhaps in riverside settlements one might argue that certain anomalies are boat quays owing to their form, but also because of their adjacency to rivers. Flour and saw mills of the historic period typically co-exist with millponds; defining the latter could help identify the former.

4.5.3.6. Experience as a basis for pattern recognition

When remote sensing projects are first initiated in a region, many anomalies may be unambiguously cultural in origin, as determined through the foregoing recognition elements. Yet, a great amount of uncertainty may remain concerning the identity of other anomalies. In these contexts one must turn to archaeological excavations as a means of anomaly identification (Section 4.5.1). An immediate source is *already published*

archaeological reports that illustrate plans of dwellings, public buildings, storage facilities, and other features. It was this tactic, in fact, that led to initial hypotheses that anomalies in Figure 4.31d represented the central hearth and roof support posts.

Beyond published literature one may also conduct limited archaeological excavations over a *sample* of anomalies of a given type and form. If consistent identifications are made then other anomalies of the same type and form likely represent the same cultural feature. For example, at Huff village, a Native American settlement in North Dakota dating to the mid-fifteenth century, Ahler and Kvamme (2000) describe hundreds of roughly circular anomalies about 1-1.5 m in diameter, detected by magnetometry (Figure 4.33a). A systematic soil coring program over a representative area, using a one-inch (2.54 cm) corer, allowed relatively secure identifications. About 74 percent were subterranean storage pits, 16 percent were hearths, and the remaining ones were small middens (Kvamme 2003). Moreover, hearths yielded characteristically higher measurements (Bales and Kvamme 2004) and were usually located along house centerlines (the association property). Storage pits, on the other hand, were often located under house entryways, off-centerline within houses, and surrounding house perimeters. This knowledge allowed remaining anomalies of this size and type to be classified with a high degree of confidence (Figure 4.33b).

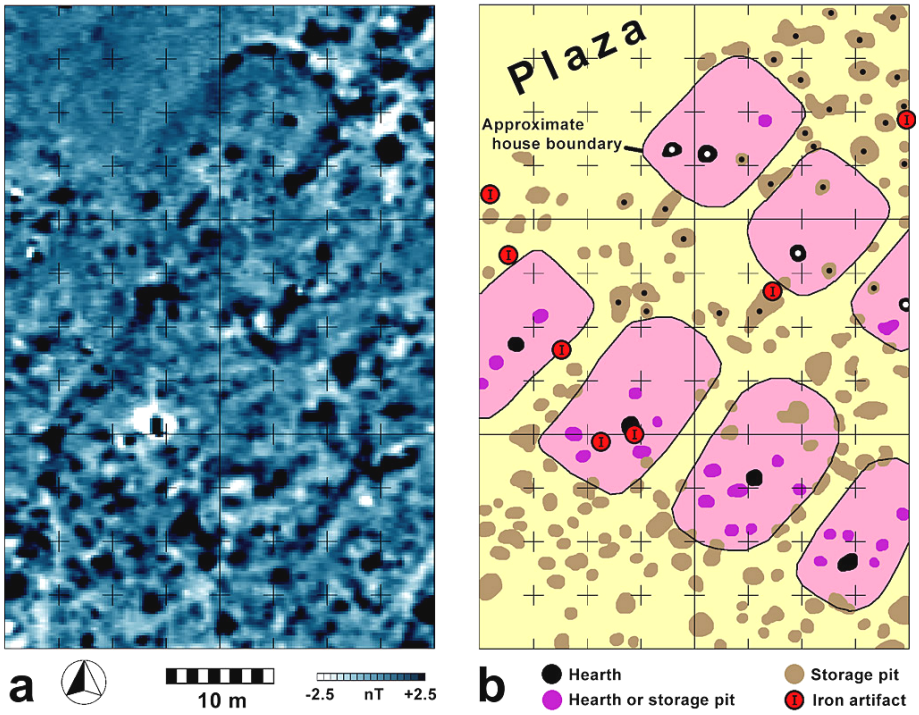


Figure 4.33. *Interpreting anomalies based on prior archaeological experience. a) Small, 1-1.5 m, circular magnetic anomalies at the 15th century Huff Village, North Dakota. b) Interpretation of significant anomalies in (a) based on a limited coring program (conducted at anomalies containing small dots) that revealed most were storage pits or hearths. Other geophysical data helped to define house outlines. Several iron artifacts are also revealed by characteristic signatures.*

By testing samples of anomalies of a particular type, size, and form, one can potentially develop an "anomaly library" for a region against which future work can refer. For example, anomalies with a characteristic range of measurements, sizes, and shapes might point to specific kinds of archaeological features. Anomaly form, size, and type will vary, of course, according to the archaeological culture, period, remote sensing technique, soil, and environmental conditions at the time of data acquisition. On-line databases of remote sensing results, such as the *English Heritage Geophysics Database* (http://www.eng-h.gov.uk/SDB) and the *North American Database of Archaeological Geophysics* (http://www.cast.uark.edu/nadag), are making libraries of results available.

It is emphasized that as one gains experience in a region with (1) the forms of archaeological expressions that occur (e.g., shapes of houses, hearths, storage pits, public buildings, defensive works, graves, etc.), and (2) with the nature of soil types and environmental conditions that influence how anomalies are formed, one's ability to generally identify anomalies increases.

4.5.3.7. Pattern recognition approaches: summary

The foregoing anomaly recognition domains are not independent of each other or of inference strategies based on knowledge of physical principles (Section 4.5.4). As a rule, we might state, for example, that a hearth goes with a house (association), the hearth is smaller than a house (size), each might be recognized by circular forms (shape), and postholes might be regularly arranged (repetitive patterns). At the same time, the physics of magnetometry suggest that hearths should yield large magnitude magnetic measurements compared to fired postholes, while burned house walls might exhibit moderate values (see below). These inferences, in fact, were all combined to form magnetic interpretations of the house at Whistling Elk village (Figure 4.31d), prior to confirmation by excavation (Toom and Kvamme 2002).

Pattern recognition is not a foolproof method. Complex archaeological deposits may make anomaly identification difficult. Intensive occupations with dense cultural stratigraphy or superimposed constructions can "jumble" the signals that might be remotely detected, making patterns unclear. Even in the most interpretable sites there are always anomalies that cannot be explained with high certainty. Moreover, every site will also contain anomalies of biological, pedological, geological, or recent cultural origin that can look like potential archaeological ones of interest. In other words, a small percentage of the anomalies classified as storage pits in Figure 4.33b might actually represent recent badger dens, for example. Recent cultural practices introduce anomalies that are particularly numerous. Something as mundane as mow marks and as silly as tethered grazing goats can generate linear or circular anomalies, respectively, in remotely sensed imagery. Former field boundaries, recent roads, trails, cattle tracks, and pipelines all can be linear in form and point to anomalies that might be prone to misinterpretation as features of possible interest. Good background research and on-site field visits are essential to reducing potential errors of interpretation.

4.5.4. Anomaly Identification through Theory

A third approach that may explain the source of an anomaly derives from knowledge of *fundamental physical principles*—the various theories and physical laws that underlie each remote sensing method. For example, stone or brick is highly resistant

compared to surrounding soils, producing large measurements of electrical resistivity; consequently, high resistivity anomalies may point to stone or brick (Gaffney and Gater 2003). Dense materials immediately below the surface (stone, a packed prehistoric house floor) tend to retain the day's thermal energy and thus appear warm to a thermal infrared radiometer in an evening image, so positive thermal anomalies may point to dense materials on or immediately below the surface (Avery and Berlin 1992). Given the large number of remote sensing detection methods, such an example list could continue indefinitely. The theoretical basis of remote sensing and variation inherent to the archaeological record is too rich and enormous to set forth in any kind of detail. The following sections therefore draw from a single method, magnetometry, to illustrate the kinds of basic principles that lead to anomaly identification (see Kvamme 2006a for a more complete discussion).

4.5.4.1. Natural causes of environmental magnetism

Three natural processes are primarily responsible for magnetic variations in near-surface deposits. They include:

1. *differences in magnetic susceptibilities* between various materials, deposits, and soils;
2. *magnetic enrichment of topsoils* stemming from physical and chemical processes that include weathering, and biogenic processes that include magnetotactic bacteria; and
3. *firing effects on soil magnetism* that produces a thermoremanent effect (Evans and Heller 2003).

Yet, people live within their environment, generally on topsoil, and modify it and other deposits. Various human behaviors therefore interact with these natural processes in ways that produce predictable, culturally induced, magnetic variations common to archaeological sites.

4.5.4.2. Anthropogenic causes of magnetic variation

Seven cultural processes can be defined that result from the interaction between human behaviors and the foregoing natural processes (Clark 2000; Scollar et al. 1990; Weymouth 1986). Knowledge of these processes and the principals upon which they are based facilitate the interpretation of magnetic anomalies.

1. *Fires* are common in human occupations. Cooking fires occur in hearths (Figure 4.31d, 4.33), fireplaces, and ovens within and outside of structures. Certain technologies require intensive firing—for ceramics, glass, or metallurgy—usually in specialized kilns or furnaces. Fires also occur accidentally when a structure burns (Figure 4.31d), or purposely when a village is razed in warfare. In all cases high temperatures introduce thermoremanent anomalies to the archaeological record (see Section 4.3.1). *Implications for interpretation:* extremely high and concentrated magnetic anomalies may point to areas of intense heating by fire. Broad anomalies may indicate a burned house or structure. Concentrated anomalies of high magnitude may be dipolar, a characteristic signature of hearths; similar anomalies broader in area and magnitude may point to kilns.
2. *Constructions and artifacts of fired materials*. Bricks are made of moderately fired clay, which increases their magnetic susceptibilities. They were common

historic period construction materials employed in large numbers. Prominent anomalies therefore result owing to their large cumulative mass (Figure 4.34). *Implications for interpretation:* a rectangular or linear anomaly might suggest a foundation or wall by its pattern. Marked positive measurements can lead to interpretations of brick walls, foundations, or pavements.

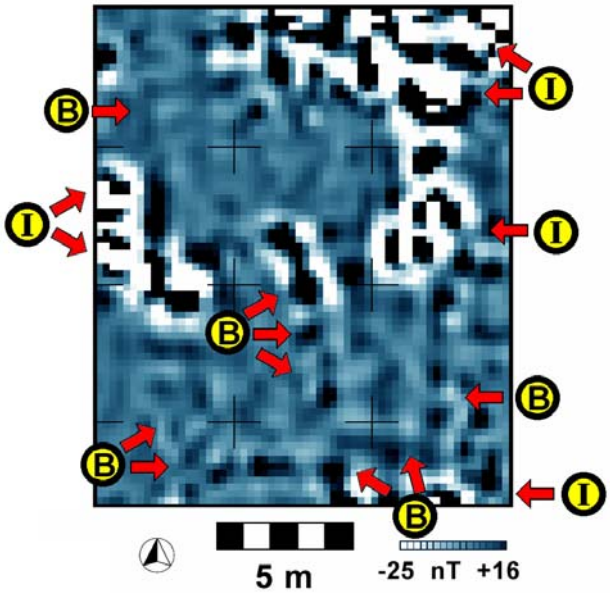


Figure 4.34. *The fired brick foundation of the Mount Comfort Church (1840s-1863), Arkansas, is visible among numerous iron-produced anomalies from stove parts, nails, and other artifacts about 35 cm below the surface, as revealed by excavation (Key: B=brick; I=iron).*

3. <u>*Magnetic enrichment of settlement soils*</u>. Fired materials within settlements become dispersed through time owing to hearth cleanings, subsequent constructions that redistribute materials from earlier hearths or burned structures, and the breakage of fired artifacts (ceramics, brick). Additionally, occupations introduce organic materials to the topsoil (food, waste products) that promote bacterial growth, including magnetotactic and other bacteria that concentrate magnetic compounds (Fassbinder et al. 1990). *Implications for interpretation:* areas with raised levels of magnetic susceptibility may point to regions of more intensive occupation or activity.
4. <u>*Accumulations of topsoil*</u>. Topsoil is usually magnetically enriched owing to several natural processes. Constructed features built of topsoil, or excavated areas filled with topsoil, cause local increases in the magnetic field. This effect is exacerbated in constructions formed of magnetically enriched settlement soils (Principle 3). *Implications for interpretation:* mildly high magnetic measurement in and about former houses may point to earth or sod constructions (e.g., in the Great Plains, Native American earthlodges or Euroamerican "soddies"). High measurements over earthen mounds and raised earthworks may indicate topsoil compositions (Figure 4.35a). Linear anomalies of high magnetism may represent topsoil-filled ditches (Figure 4.35b). Small-diameter anomalies showing moderately high readings can mean topsoil-filled storage pits (Figure 4.33).

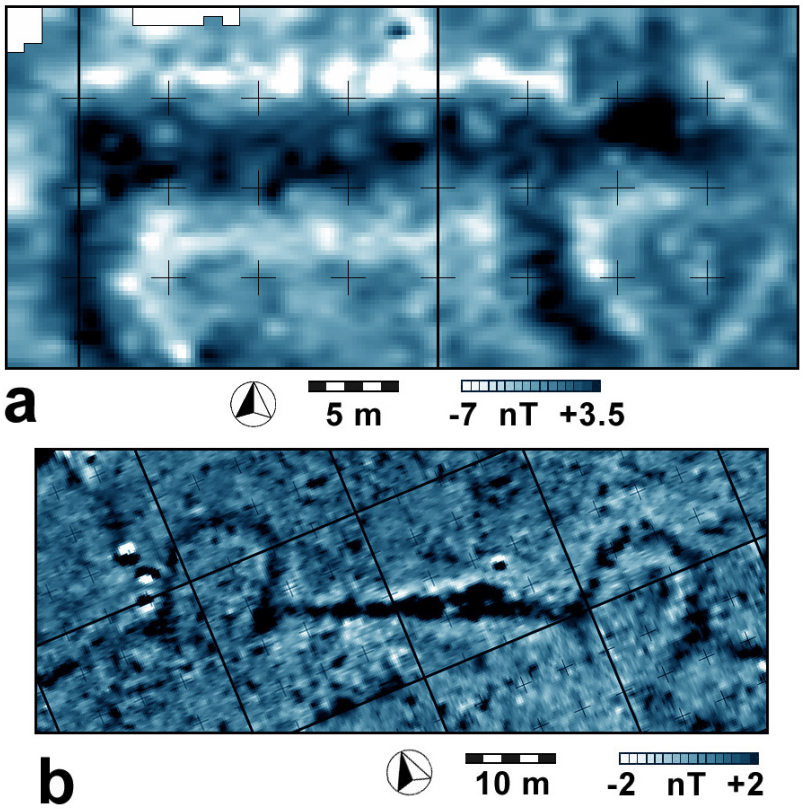


Figure 4.35. *Accumulations of magnetically enriched topsoil generate positive anomalies. a) Topsoil mounded to a height of about 1 m forming the Great Bear effigy (ca AD 650-1300) at Effigy Mounds National Monument, Iowa (source: Kvamme 1999). b) Surface soils completely fill a fortification ditch and bastions, obscuring them visually but not magnetically, at the Double Ditch State Historic Site (32BL8), an ancestral Mandan village in North Dakota (ca AD 1500-1780; source: Kvamme 2003).*

5. <u>*Topsoil removal.*</u> Some constructions remove topsoil, resulting in a net lowering of the magnetic field or "negative" anomalies (there is less magnetic material than in neighboring areas). Unfilled ditches (Figure 4.36a), recessed house floors, subterranean storage pits, cellars, and even former archaeological excavations and looter's "pot holes" effectively remove small to large volumes of magnetically enriched topsoil, causing negative contrasts (Figure 4.36b). Simple incisions in the ground caused by foot traffic along trails or roads (Figure 4.36b) can produce the same effect, where topsoil has been kicked or pushed aside. Occasionally, sediments and soils removed during the excavation of a grave are not replaced in their original order. If the more magnetic topsoil does not lie at the surface after a grave is filled negative anomalies can sometimes be expressed (Figure 4.36c). *Implications for interpretation:* anomalies indicated by negative measurements indicate something about construction methods, whether topsoil was removed, deeply buried, or whether a void might exist below the surface.

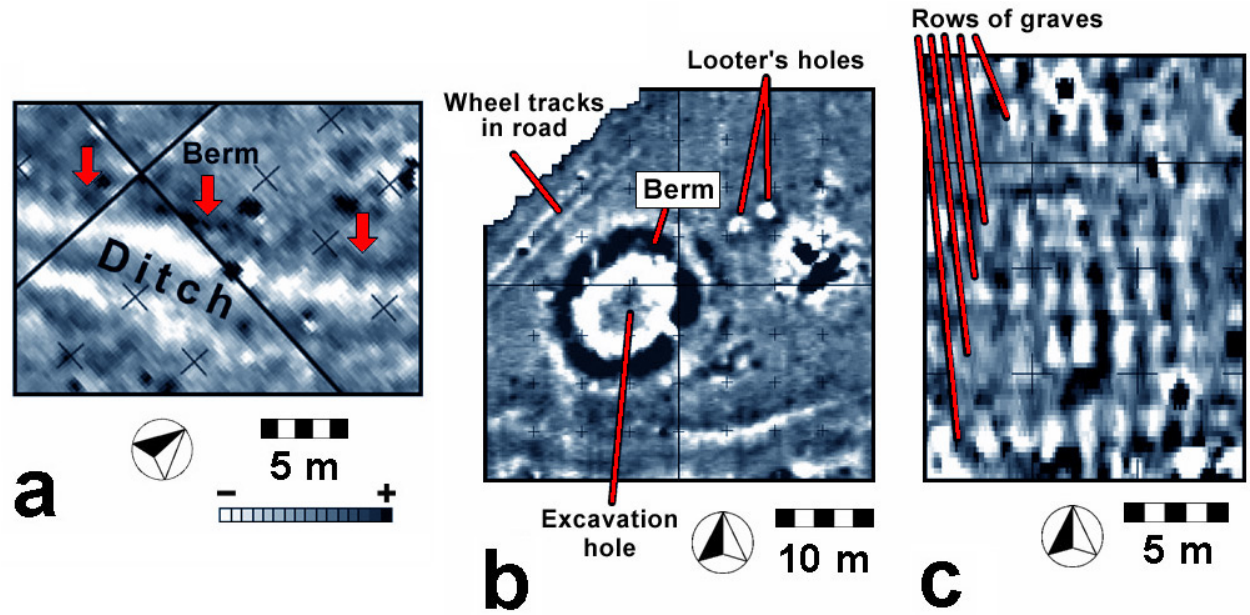


Figure 4.36. *The removal of magnetically enriched topsoil can produce negative anomalies. a) At the Menoken Village State Historic Site (32BL2), a fortified village (ca AD 1240) in North Dakota, negative anomalies are seen over the slopes of a mostly unfilled fortification ditch (the bottom of the ditch is somewhat magnetic owing to its partial in-filling with eroded topsoil or pedogenesis; source: Kvamme 2003). (A raised berm adjacent to the ditch indicates positive anomalies.) b) Incised vehicle tracks, an unfilled 1930s archaeological excavation, and looter's pits form negative magnetic anomalies at Menoken Village. c) At the Confederate Cemetery (founded in 1862) at the University of Mississippi, numerous interments are indicated as negative anomalies, most likely because topsoil was not replaced at the top of each grave (source: courtesy of Jay Johnson, University of Mississippi).*

6. *Import stone and other materials.* Building foundations, pavements, and floors are often constructed of stone. Many rocks tend to be more magnetic (e.g., igneous rocks), but some are less magnetic (e.g., most limestones) than surrounding soils, thereby generating magnetic contrasts (Figure 4.31c). Likewise, specialized sediments are sometimes imported for desirable characteristics, such as clay for prepared floors or pit linings, sand or gravel for trails, walkways, or base substrates for larger constructions. These materials, too, can produce detectable contrasts depending on their inherent magnetic susceptibilities, volume, and depth. *Implications for interpretation:* A rectangular or linear anomaly might suggest the presence of a foundation, pavement, or wall. Positive or negative magnetic values can lead to inferences concerning material types if common building materials in a region are known.
7. *Iron artifacts*. Iron artifacts tend to be readily detected by magnetometry depending on their size, shape, orientation, mass, and depth below the surface. Iron is nearly ubiquitous in historic period sites and markedly alters the earth's magnetic field, commonly producing large and easily recognized dipolar anomalies (Figure 4.33-4.34). *Implications for interpretation:* The magnetic signature of iron artifacts is usually unambiguous owing to its concentrated, large magnitude, and dipolar form, leading immediately to an inference of ferrous

metal. In historic sites iron litter may point to significant artifacts, dumping areas, or the loci of former wooden structures that employed nails in their construction.

4.5.5. Discussion

As with pattern recognition tactics, inferences based on physical principles are not made in isolation. Multiple sensing-method surveys respond to many physical properties of the archaeological record and thus can offer several lines of evidence. A broad linear feature revealed by high magnetic measurements could be classified as a possible burned feature, filled ditch, or a brick or stone pavement, for example. The complementary information of a resistivity survey could improve the probability of correct identification. Moderately high resistivity might point to brick, very high measurements might suggest stone, low values might indicate a sediment-filled ditch, and neutral values might support the hypothesis of a burned feature. If a shallow construction, and surveyed by ground-penetrating radar with a high frequency antenna, individual stones or bricks of a pavement might be resolvable; a mid-frequency antenna could outline the cross-section of a buried ditch.

4.6. GEOPHYSICAL RESULTS

The following sections summarize the ground-based geophysical findings at each of the project sites. The archaeological nature and culture history of each site was described in Section 3.3. The goal here is to summarize the geophysical surveys and methods employed at each project site and to suggest in a broad and general way what those findings indicate about the sites. In other words, using the interpretational principles of Section 4.5, some of the more obvious anomalies are explained here. This material is offered as background in order that some understanding of results is reached prior to data integrations discussed below. Full interpretation of the data is examined after more complete examination of the data through their integration.

4.6.1. Geophysical Results at Army City

Kenneth L. Kvamme, University of Arkansas

Five ground-based geophysical methods were employed at Army City, more than any other of the project sites. Each covered the entire study block measuring 100 x 160 m (1.6 ha) and centered over the town's former commercial core (see Figure 3.2). Prior geophysical work by Hargrave et al. (2002) showed good geophysical results using electrical resistivity and magnetic gradiometry surveys, with the former responding to highly resistant concrete walls and floors and the latter to the many ferrous metal artifacts endemic to late historic sites and areas of burning from the 1920 fire (Section 3.3.4.1). Test excavations (Hargrave 1999b; Kreisa and Walz 1997) and the period of occupation suggested the likelihood of a network of underground pipes for water or sewage (the earlier magnetic survey, which would have confirmed this, did not cover the core village area). Although EM quadrature phase data are theoretically the inverse of soil resistivity—measuring *electrical conductivity*—it was thought that EM methods could potentially reveal new information because metal artifacts yield pronounced anomalies that do not appear in probe-contact resistivity surveys (Bevan, 1998:36-39). Moreover, the great fire of 1920 quite likely raised magnetic susceptibility in the near surface, so the

in-phase component of an EM survey, limited to very shallow depths, was likely to prove insightful. Despite the assertion by Hargrave et al. (2002:94) that "the high clay content of the soil at Army City eliminated GPR as an effective technique" (clay is generally conductive and disperses radar energy), prior experience, a long period of drought (dry soils lower conductivity), new software for signal processing, and other studies suggested the possibility of obtaining positive results. A full GPR survey of the study area was therefore pursued because this technique can measure yet a different physical property of the sub-surface, dielectric contrasts, and it offered the possibility of three-dimensional imaging at a variety of depths. Despite the destruction of the town by fire, wholesale movement of buildings to the nearby town of Ogden, and the dismantling of the remainder such that few hints remain on the surface, Hargrave et al. (2002) demonstrated significant subsurface archaeological content yet remains through a program of geophysics and test excavations. Concrete and masonry foundations and floors immediately beneath the surface, pipes and pipe trenches, large areas of burning, streets and street gutters, the possibility of cellars, all recommended the likelihood of excellent geophysical findings.

Table 4.1. *Geophysical methods, instruments, sampling densities, and principal anomaly types at Army City.*

Geophysical Method	Instrument	Original resolu-tion (m)	Prospec-tion depth (m)	Unit of measurement	Principal indicated anomalies
Conductivity	Geonics EM-38	.5 x 1	1.5	mS/m	Metal pipes
Ground-penetrating radar	GSSI SIR 2000 & 400 MHz antenna	.025 x .5	2.3 (60 ns time window)	amplitude (16-bit scale)	Foundations, floors, street gutters
Magnetic gradiometry	Geoscan Research FM-36	.25 x 1	1.5	nT/m	Iron pipes, large iron artifacts, burned areas
Magnetic susceptibility	Geonics EM-38	.5 x 1	< .5	ppt (received to transmitted signal)	Burned areas, foundations, street gutters, iron pipes
Electrical resistivity	Geoscan Research RM-15 twin-probe array	.5 x 1	.5	ohm/m	Concrete floors, foundations, building footers, gutters

The five ground-based surveys of electrical resistivity, magnetic gradiometry, ground-penetrating radar, quadrature-phase EM for soil conductivity, and in-phase EM for magnetic susceptibility, were undertaken during a two-week period in the summer of 2002. All except GPR were conducted within survey blocks of 20 x 20 m, but the GPR survey was conducted in larger blocks of 50 x 50 m or 30 x 50 m to avoid numerous field set-ups and data processing difficulties associated with concatenation of many small units (these data were employed to illustrate geophysical data processing sequences in Section 4.4.5). Instrumentation, original spatial resolutions, and approximate prospecting depths

for each survey are summarized in Table 4.1. All the data were resampled to a uniform spatial resolution of .5 x .5 m for subsequent data integrations.

The five geophysical data sets are illustrated in Figure 4.37. They collectively indicate various aspects of Army City's commercial core that was centered on Washington Avenue, down which the historic 1918 photograph of Figure 3.2 was shot. That large southeast-northwest trending avenue is well revealed by the geophysical results, particularly by magnetic susceptibility in Figure 4.37e, as are other side streets (see Figure 3.1 for a plan map of the town's streets). Foundations and floors of many of the town's buildings are also seen, as is the infrastructure of sewer and water pipelines (Figure 4.37).

4.6.1.1. Anomalies common to all data sets

All of the geophysical data in Figure 4.37 robustly or mildly point to former structures, including floor areas, walls, and debris fields, and all at least hint at the former street grid, either by showing their surfaces or defining their edges (gutters or curbs).

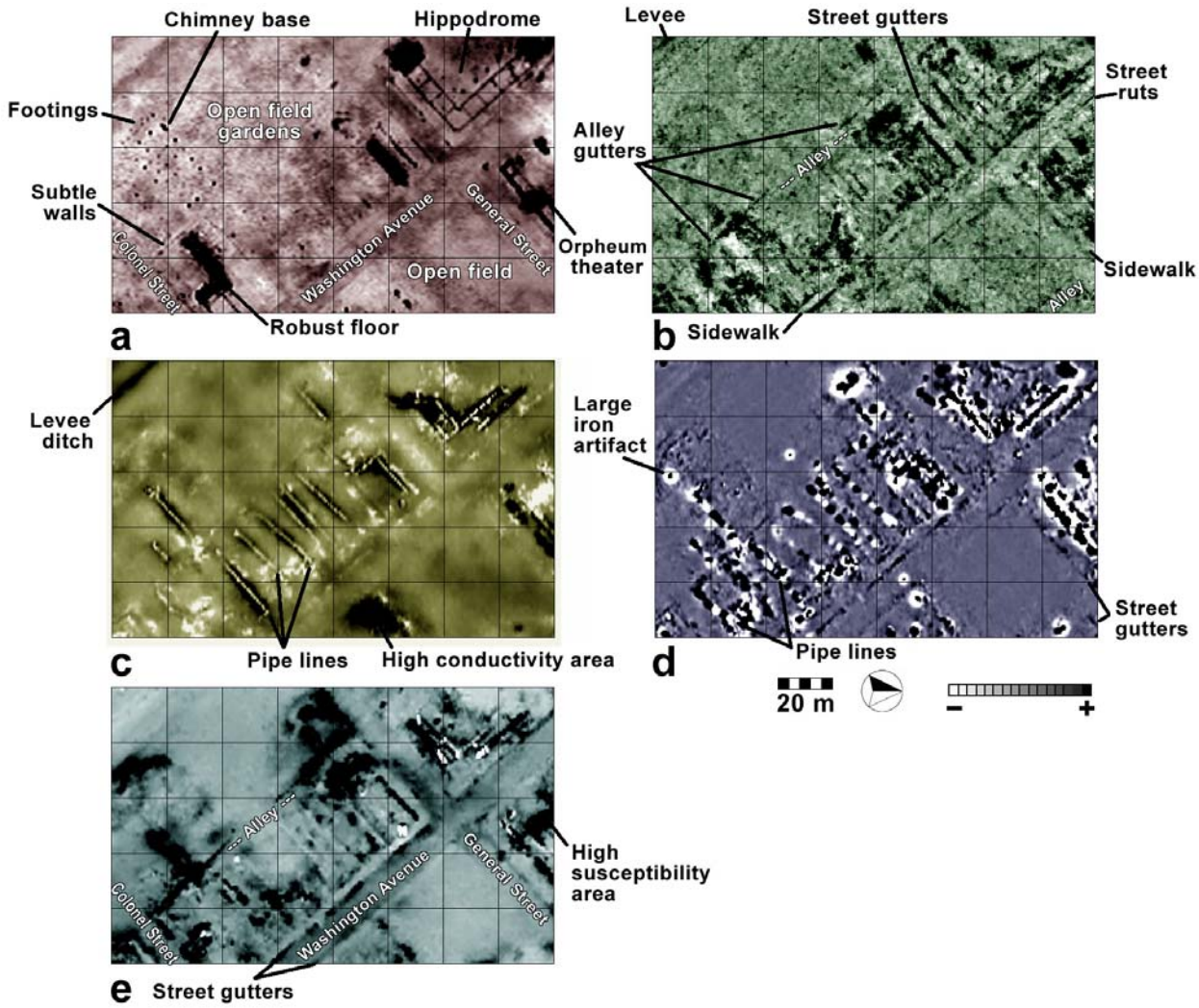


Figure 4.37. *Interpretations of Army City geophysical data: a) electrical resistivity, b) ground-penetrating radar, c) EM conductivity, d) magnetic gradiometry, e) magnetic susceptibility.*

All data sets point to two broad zones with a general absence of anomalies The larger one to the southwest was an open space partially used for gardens (see historic photo in Figure 3.2), while the smaller one to the northeast was an open field during the occupation (as revealed by historical photographs in Rion 1960). A modern drainage canal and a portion of a levee are also visible in each data set (Figure 4.37).

4.6.1.2. Electrical resistivity

The Army City electrical resistivity data is particularly clear in its portrayal of certain structures (Figure 4.37a). Several broad areas of high resistivity point to floors of likely concrete or masonry. The extremely robust foundation walls of the Hippodrome are well indicated, suggesting concrete or brick, as are several other prominent foundation walls of other structures. Less robust foundation walls of buildings are also indicated, particularly in the southern part of the study block. These may indicate narrower walls of masonry or concrete, or perhaps only builder's trenches that once held walls within. Of particular interest are a series of anomalies that represent footings showing the outline of a former structure to the southwest, seen in the historic photo of Figure 3.2. The largest oval shaped anomaly probably represents the base of a chimney, seen in that photo. Other more massive footing or piers occur within the Hippodrome. The high resistivity of much of General Street west of Washington Avenue suggests it may have been paved or graveled with a resistant material.

4.6.1.3. GPR

This GPR data set is a depth-slice showing maximum reflection amplitudes from 20-40 cm below the surface (Figure 4.37b). It complements the resistivity data because it responds to many more of the town's foundation walls and floors. Many walls are robustly indicated, as are numerous subtle walls. Street gutters, in particular, are clearly seen suggesting significant material changes in their compositions. This is also true of a series of gutters that apparently parallel alleys. It is also of interest that the front lines of buildings on Washington Avenue and General Streets are clearly shown with a narrow space separating them from the street gutters. This marks the area of sidewalks that are known to have been present, as seen in the historic photograph of Figure 3.2. Within Washington Avenue several linear features may be seen that may point to street ruts or "puddle" areas (also seen in the historic photo).

4.6.1.4. EM conductivity

EM conductivity, highly sensitive to metals of any kind, primarily reveals Army City's water and sewer lines that lie beneath the remains of some of its principal buildings (Figure 4.37c). They are indicated because metals are highly conductive, but the massive amounts of metal they represent cause EM instruments to become "saturated" and yield negative anomalies (Bevan 1998). Several "halos" of moderately low conductivity suggest the outlines of buildings—probably debris fields of resistant building materials (bricks, mortar, concrete) dispersed from their destruction or dismantling. General Street, west of Washington Avenue, exhibits low conductivity—paralleling the resistivity findings—suggesting it might be surfaced with resistant material. Several street edges show the same phenomenon, perhaps indicating resistant gutter fill or the presence of concrete curbs (Army City is known to have had concrete

sidewalks, Hargrave et al. 2000). Two areas illustrate high conductivity. One is associated with the Hippodrome and the other with a likely structure on the east side of Washington Avenue. They may indicate the presence of moist, conductive, fill sediments within rooms of those structures, or possibly filled cellars.

4.6.1.5. Magnetic gradiometry

The magnetic gradiometry data are particularly responsive to ferrous metals (Figure 4.37d). Although de-spiking was employed to remove the plethora of small point anomalies from the sea of iron artifacts (mostly nails) covering the site (Section 4.4.1.1), larger ones remain that point to massive iron artifacts. Included among these are the water and sewer lines also revealed by EM conductivity. Unlike EM conductivity, pipes are represented as dipolar anomalies, with each pipe joint yielding a positive and negative pole. Certain areas of former structures also illustrate large anomalies pointing to ferrous metals, possibly rebar and steel used in their construction. Broad areas of high magnetism may point to fired soils associated with the fire of 1921. That street gutters are visible may point to fills composed of materials of high magnetic susceptibility.

4.6.1.6. Magnetic susceptibility

The in-phase EM survey yields data proportional to magnetic susceptibility, and a very different view of Army City (Figure 4.37e). Streets and gutters of streets and alleys are particularly prominent supporting the idea of fill sediments of high magnetic susceptibility. They may derive from natural topsoil transported by wind or rain, and quite possibly from burned materials from the town fire. Broad areas of high susceptibility may point to areas of fired soil or perhaps constructions of fired brick. Some shallow metal artifacts and pipes also are apparent in these data owing to their high susceptibility.

4.6.2. Geophysical Results at Pueblo Escondido

Eileen G. Ernenwein, University of Arkansas

Geophysical data were collected at Pueblo Escondido on three occasions. The first visit to the site occurred during a two-day reconnaissance in May 2003. It revealed abundant clusters of lithic and pottery debris and fire-cracked rock in a very large open area. In addition, small, subtle mounds were witnessed that appeared to indicate melted adobe or other remains of ruined house walls that were observed in good state during the 1960s (see Section 3.3.4.2). Some of these mounds were probably caused by accumulations of wind-blown soil around the bases of bushes that have since died. Others appeared in rectangular patterns, however, suggesting a cultural origin and possible adobe remnants. Based on these surface features, two areas, separated by a small access road, were chosen for survey (labeled “A” and “B” in Figure 3.4a). The geophysical surveys included 4,000 m^2 by electrical resistivity and magnetic gradiometry, and 1,600 m^2 by ground-penetrating radar and electromagnetic induction methods (the last yielding conductivity and magnetic susceptibility data). These results were not very insightful, however, with few anomalies of clear cultural origin.

The same instruments were brought to the site the following October, when it was hoped it would be cooler and usual fall rainfall would yield moist soil conditions for

electrical resistivity data capture. A 400 MHz GPR antenna was employed instead of the 900 MHz antenna used the previous spring because it would allow deeper penetration and reduction of noise common to the high frequency antenna. Additionally, the area of investigation was moved southward in the hope of hitting significant site elements, since few were indicated in the previous survey areas. Despite this planning, temperatures were very warm and the soil was very dry, precluding collection of resistivity data. Yet, the new study area, a full hectare (10,000 m^2), did encompass numerous significant cultural features and the lower-frequency GPR antenna successfully detected and mapped them (area "C" in Figure 3.4a). The data acquired in these surveys were of exceptional content and quality, and were used during the subsequent year for data integration experiments and planning of excavations for the field validation phase of the project (Section 3.3.4.2). The final trip to the site was for the anomaly validation phase. Advantage was taken of this trip to expand the irregularly shaped study area to form a full rectangle, with geophysical surveys conducted side-by-side with excavations in an additional 2,000 square meters. This study area is labeled "D" in Figure 3.4a. Instrumentation, original spatial resolutions, and approximate prospecting depths for each survey are summarized in Table 4.2.

Table 4.2. *Geophysical methods, instruments, sampling densities, and principal anomaly types at Pueblo Escondido.*

Geophysical Method	Instrument	Original resolu-tion (m)	Prospec-tion depth (m)	Unit of measurement	Principal indicated anomalies
Conductivity	Geonics EM-38B	.25 x .5	1.5	mS/m	Broad zones of high conductivity-houses, melted adobe, vegetation patterns?
Ground-penetrating radar	GSSI SIR 2000, 900 & 400 MHz antennas	.025 x .5	1.6 (60 ns time window)	amplitude (16-bit scale)	House walls, floors, small pits, walls
Magnetic gradiometry	Geoscan Research FM-36	.25 x .5	1.5	nT/m	Burned walls, hearths, igneous rock alignments?
Magnetic susceptibility	Geonics EM-38B	.25 x .5	< .5	ppt (received to transmitted signal)	Burned areas, walls, burned walls
Electrical resistivity	Geoscan Research RM-15 twin-probe array	.5 x .5	.5	ohm/m	Broad zones of high conductivity-houses, melted adobe, vegetation patterns?

Taken together, the geophysical data gathered at Pueblo Escondido—up to 1.69 ha—reveal astonishing detail about the site, including the size, shape, and locations of several pithouses and pueblo room blocks, numerous other features, as well as its overall layout and organization.

4.6.2.1. Magnetic Methods

Two different magnetic data sets were acquired at Pueblo Escondido: (1) magnetometry with a Geoscan Research FM-36 fluxgate gradiometer and (2) magnetic susceptibility using the in-phase component from a Geonics EM-38 soil conductivity meter. Both methods were successful, although the results are strikingly dissimilar. Magnetic surveys are seldom conducted in desert environments because it is thought the lack of soil development precludes the formation of strong anomalies through past manipulations of the topsoil (e.g., by mounding in constructed features; see Section 4.1.4). The resulting paucity of strong anomalies in the data was complicated by the fact that wind deposition tends to concentration magnetic minerals in lag deposits around the bases of the scrubs and brush that dot the desert, creating many small anomalies because the mineral accumulations, although weak, are on the surface and therefore closer to the sensors than subsurface archaeological remains. Both contributed to a data set that only moderately reveals the subsurface, but it does include several strong anomalies illustrating geometric patterns of clear cultural origin (Figure 4.38a). Many of these rectangular and linear features (arrows) suggest the presence of several habitation structures that were either burned or possess magnetic materials (such as igneous rock) incorporated in their construction. The very robust anomaly in the northwest corner of Figure 4.38a was found during the last hours of survey, prompting a last-minute expansion of the study area to incorporate two additional 20 x 20 m squares to the west and one 10 x 10 m unit to the north. The many smaller anomalies distributed throughout the region could be hearths, roasting pits, or storage pits, but the identification of small amorphous anomalies is very difficult before excavations are conducted because many probably represent only wind-blown mound deposits beneath standing brush.

The acquisition of high-density magnetic susceptibility data by in-phase electromagnetic induction survey is rarely undertaken in archaeology because its utility is not widely known and its limited prospecting depth (less than a half-meter) makes it of limited usefulness. It probably has never been employed in a desert environment. This survey at Pueblo Escondido revealed a widespread distribution of long linear anomalies trending in the cardinal directions, as well as some large rectangular to ovoid features suggestive of pithouses (arrows, Figure 4.38b). The small linear and rectangular anomalies probably represent house walls, while the larger amorphous ones could point to earlier pithouses. The long linear anomalies surrounding some of these features are more difficult to interpret, but they could represent patio areas or *ramadas*.

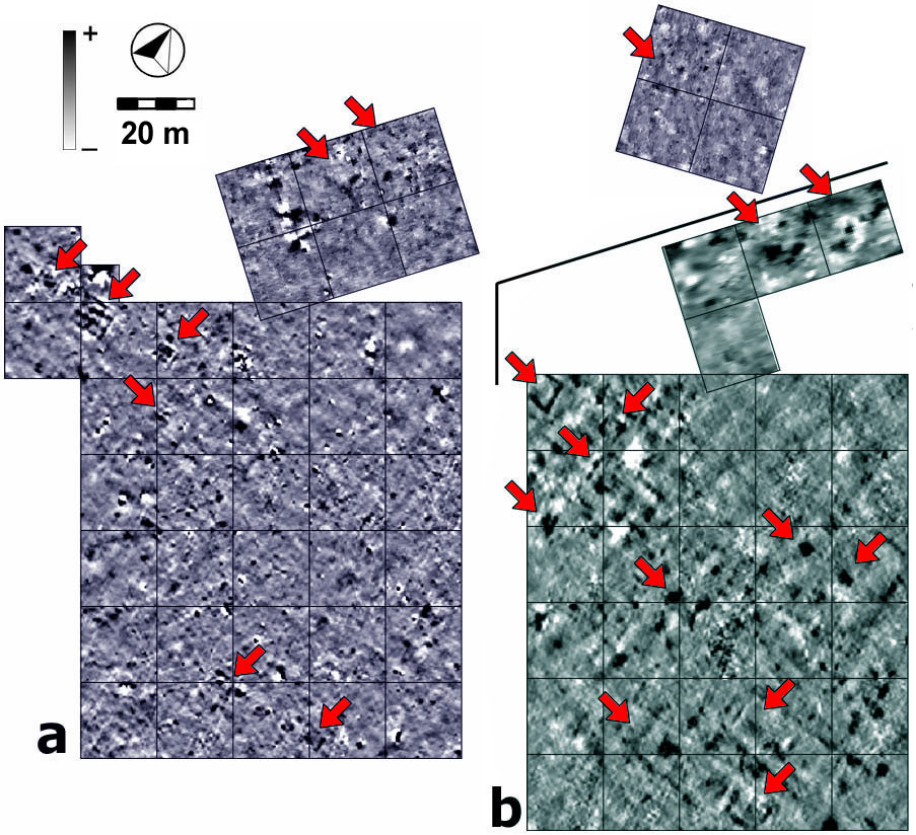


Figure 4.38. *Geophysical survey results at Pueblo Escondido. a) Magnetic gradiometry. b) Magnetic susceptibility through in-phase electromagnetic induction survey.*

4.6.2.2. Electrical Methods

Relatively moist ground conditions during the first visit to Pueblo Escondido permitted the collection of electrical resistivity data using a Geoscan Research RM-15 with half-meter twin probe spacing. Areas A and B (0.4 hectares) were completely surveyed (Figure 3.4a). During this time conductivity data were also acquired using a Geonics EM-38B soil conductivity meter, with which the southwestern two-thirds of Area B was surveyed. These data are strongly correlated with electrical resistivity in this site.

During the second data collection trip the ground was too dry, causing infinite probe contact resistance and an inability to collect additional resistivity data. Conductivity data were acquired, however, and the results showed a pattern similar to the earlier data. Both data sets are illustrated side-by-side in Figure 4.39a for the respective zones of survey (i.e., A-D of Figure 3.4a). Taken together, these data illustrate variation in electrical properties about the site. The broad variation indicated does not distinctly point to cultural anomalies, such as those apparent in Figure 4.38, but some of the positive anomalies (arrows) may possibly represent habitation structures (e.g., broad zones of "melted" adobe) or midden deposits. Evidence given in Section 5.16.2, however, strongly suggests these patterns of soil conductivity reflect only the surface distribution of vegetation at the time of the survey (compare Figure 3.4a, for example).

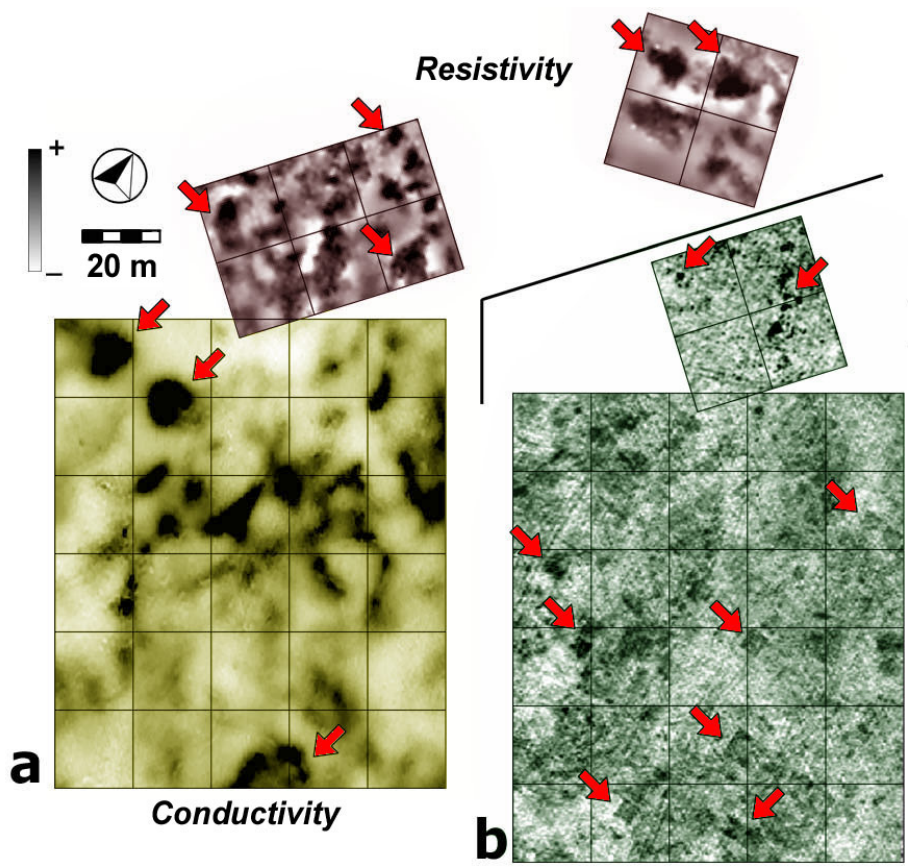


Figure 4.39. *Geophysical survey results at Pueblo Escondido. a) Electromagnetic conductivity (yellow) and electrical resistivity (red). b) GPR depth-slices representing 15-31 cmbs. The small upper block (Unit B) is from a 900 MHz antenna; the larger lower block (Units C, D) is from a 400 MHz antenna.*

4.6.2.3. Ground-penetrating Radar

The abundant evidence of surface erosion suggested that Pueblo Escondido is severely deflated, so subsurface features could be situated very close to the surface. With very shallow deposits, and dry loamy ground conditions, a GPR survey with a 900 MHz antenna was undertaken in hopes of a more detailed characterization than would be possible with the more commonly used 400 MHz antenna. This 900 MHz antenna was used with a GSSI SIR 2000 system and survey wheel during the first visit to the site, when the southwestern two thirds of Survey Area B was surveyed (Figure 3.4a). The resulting data were perhaps too detailed, revealing many irregularities and high-frequency clutter. Subsequent surveys in the main study block (Areas C, D) therefore employed a 400 MHz antenna.

Several depth-slices were generated from the GPR data. In Figures 4.39b and 4.40a,b the 900 MHz data of Survey Area B are shown in the top smaller block and the 400 MHz data of Areas C, D in the lower larger block. These data reveal that Pueblo Escondido is, indeed, very intact as an archaeological site, with significant subsurface architectural features yet remaining. The 900 MHz data do illustrate somewhat more noise and clutter, but several noteworthy anomalies are indicated (arrows). The 400 MHz data appear less cluttered and depth penetration is good. A depth-slice from 15-31 cm below surface (cmbs) reveals the upper portions of several rows of houses and associated cultural features, along with lineations caused by vehicular disturbances to the near surface (arrows, Figure 4.39b). A deeper depth-slice, spanning 31-47 cmbs, illustrates a more detailed pattern of houses (arrows, Figure 4.40a), and many are still clearly visible

in the lowest slice, 47-63 cmbs, indicating they may be present at considerable depth (Figure 4.40b). A myriad of small anomalies quite possibly indicates storage pits, hearths, roasting pits, or middens.

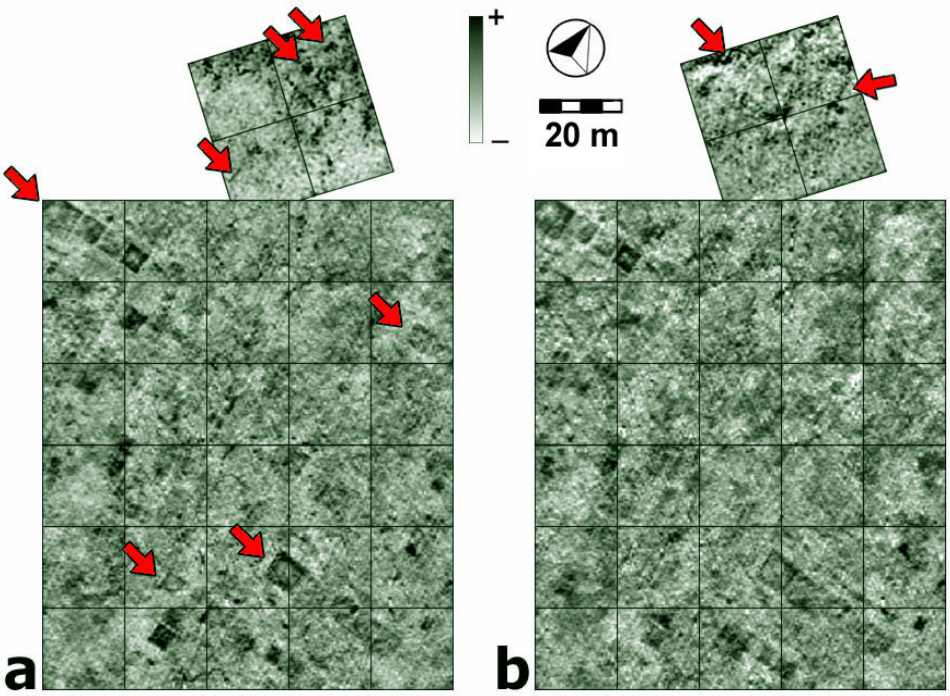


Figure 4.40. *GPR survey results at Pueblo Escondido. a) Depth-slices representing 31-47 cmbs. b) Depth-slices representing 47-63 cmbs. In each, the small upper block (Unit B) is from a 900 MHz antenna; the larger lower block (Units C, D) is from a 400 MHz antenna.*

4.6.2.4. Advanced computer-visualization of the Escondido GPR data
Michele L. Koons, University of Pennsylvania

The middle and deep slices of GPR suggest some depth to architectural features at Pueblo Escondido. Advanced visualization tools are examined here to further investigate the nature of the indicated GPR anomalies. Specifically, ways for visualizing GPR data in three dimensions (3-D) are summarized (see Koons 2005 for full details). Traditional representations of GPR data are typically presented in two-dimensional time-slice or depth-slice maps, which show cross-sections of the subsurface increasing with depth (e.g., Figure 4.40). These slices represent 3-D space, but do not demonstrate the data as a continuous three-dimensional volume. Representations of this sort have been made in the past utilizing a variety of different software and approaches. The drawback is that only one software package, *GPR Slice*, is specifically designed for GPR and archaeology. Most 3-D visualization software is made specifically for the medical industry and for geotechnical purposes; they are expensive and not completely compatible with GPR data. Three software systems are explored here to visualize the GPR data: *Slicer Dicer*, *3-D Master*, and *MATLAB 7.0*.

<u>*Data Pre-processing*</u>. Data from a single 20 m survey block are investigated (this unit is indicated by the upper left arrow in Figure 4.40a). That block reveals a series of square anomalies aligned in a row that represent adjacent pithouses or perhaps rooms in an above-ground pueblo-style room block (see Section 8.16.2 for confirmation of these

identifications). High and low pass filters were set to 200MHz and 800MHz to remove excessive noise from the traces and the time window was opened to 30 nanoseconds. Background noise was removed, which facilitated the migration of reflections back to their original sources and smoothed the raw data files, leaving only source reflections of potential significance in the profiles. The program *GPR_Process* was then used to slice the data into segments one nanosecond thick. Each one-nanosecond slice represents a layer of subsurface material about 16 cm thick. Only the upper 13 slices were used for this project, however, because few recognizable cultural features appeared in the lower slices. The pit-room structures and other linear and smaller features are apparent within the data (Figure 4.41).

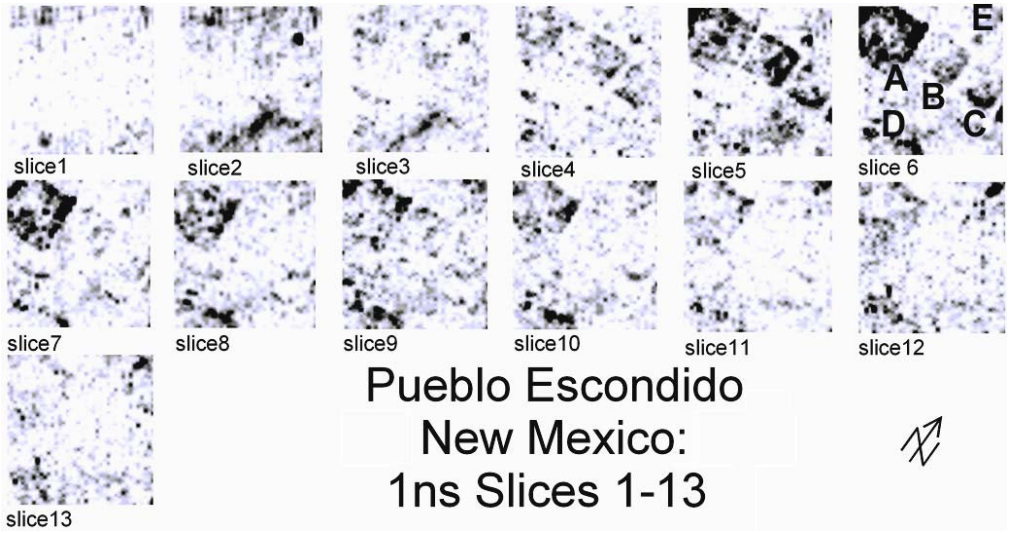


Figure 4.41. *Thirteen one-nanosecond thick time-slices, each representing approximately 16 cm of subsurface "thickness." Certain anomalies are labeled in Slice 6.*

Certain problems do arise when attempting to make a 3-D volume from the 2-D GPR slices. For example, deeper reflections will almost always have lower amplitudes than shallower reflections because of the conical shape of the GPR "footprint" and the attenuation of energy as it propagates through the ground (Conyers 2004:91). The amplitudes in each slice were therefore normalized by dividing them by the mean slice amplitude. This process effectively normalized the gains so the numbers were within a comparable range, although inconsistencies between slices remained.

Slicer Dicer. Individual time-slice maps are imported into *Slicer Dicer* as graphics files and stacked on top of one another. The transparency threshold can then be manipulated until the display appears discrete. Depending on the selected threshold, fewer or more anomalies may be defined. This is exemplified in Figure 4.42 where Slice 7 is imported, thresholded, and ultimately stacked with other thresholded slices. Figure 4.42c shows a 3-D volume rendering with liberal thresholding that defines a minimum anomaly set.

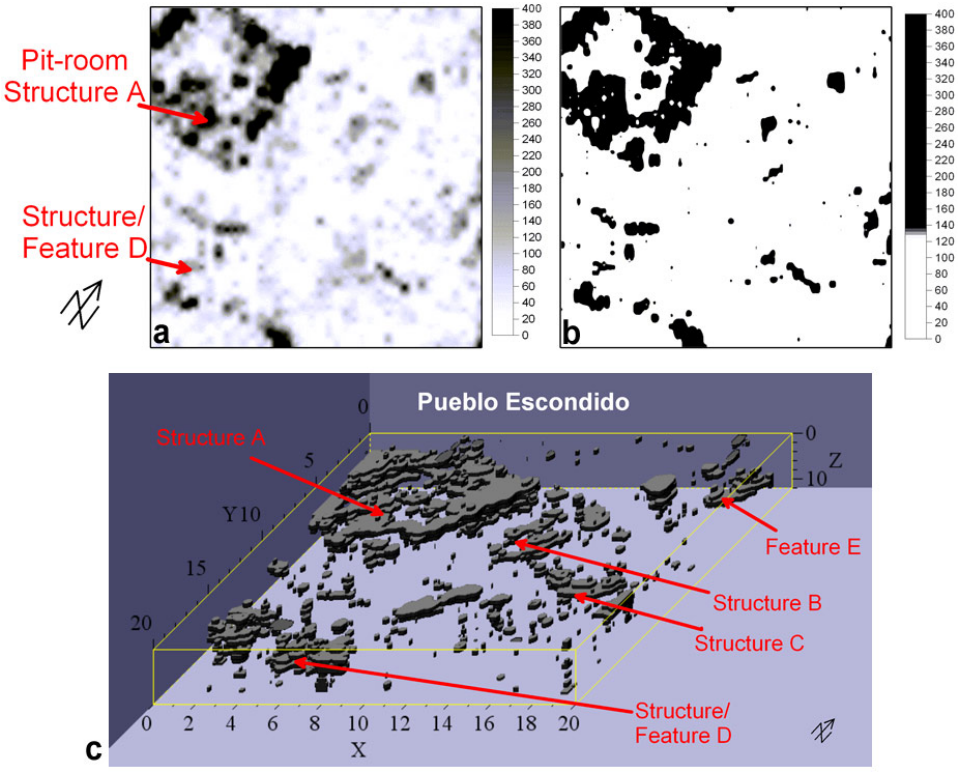


Figure 4.42. *3-D visualization with Slicer Dicer software. a) Slice 7. b) Slice 7 after thresholding to define significant anomalies. c) 3-D visualization of all 13 slices at a liberal threshold that defines only the most prominent anomalies.*

<u>*3-D Master*</u>. This software was specifically developed for the visualization and animation of ground-water flow. Unlike *Slicer Dicer*, which accepts graphics files as input, *3-D Master* requires numeric X, Y, Z and amplitude data to create surfaces. For the isosurfaces to be accurate the GPR data must be normalized between slices, as described earlier. The result, however, is problematic because portions of known features are missing (Figure 4.43). This is partially a product of the mean normalization algorithm, because anomalies with amplitudes lower than the mean are not well represented. Other forms of normalization can partially correct this problem, but noise and small isolated features still will complicate the volume.

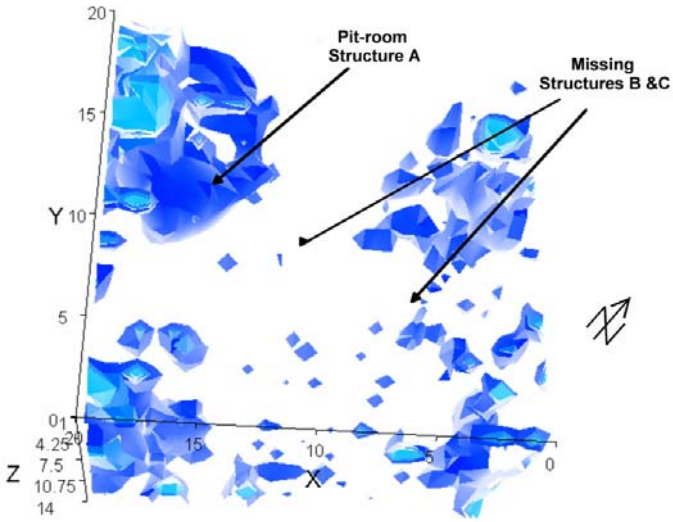


Figure 4.43. *3-D visualization of isosurfaces describing GPR anomalies with 3-D Master software.*

<u>*MATLAB 7 and the Canny Algorithm*</u>. MATLAB (which stands for Matrix Laboratory) refers to both a programming language and an environment for numerical computing by manipulating mathematical matrices. This software can be used for analyzing and visualizing data by applying preexisting functions and creating new ones for specific tasks. Canny edge detection, developed by John F. Canny in 1986, is a three-step algorithm used to locate continuous edges within an image. This algorithm is known as the optimal edge detector because the stages take into account the desire for a low error rate, optimal edge localization and minimal or only one response to a single edge (Canny 1986). The Canny edge detection operator receives focus for handling GPR data, because a goal is to define discrete archaeological entities within a continuous 3-D data volume. This algorithm works in multiple stages, but follows the three main steps of image edge detection: noise smoothing, edge enhancement and edge location ().

The normalized X, Y, Z and amplitude data from the 13 GPR slices were first combined into a single text file and imported into MATLAB. With 700 data points in the X (or transect) direction and only 40 in the Y direction in each individual slice (due to unequal GPR field sampling; see Section 4.1.7), the data had to be resampled to a uniform matrix of 40 x 40 pixels to avoid potential problems this might cause in edge detection (Figure 4.44a). The first step in the Canny algorithm applies a Gaussian-weighted low-pass or smoothing filter to each 2-D slice to reduce noise, consolidate anomalies, and preserve image edges (see Section 4.4.2.2) (Figure 4.44b). Next, a Sobel filter is applied to each 2-D slice to locate edges, with the result referred to as a "strength image" which yields high spatial gradient measurements corresponding to edges, represented by "ridges" in the result (Trucco and Verri 1998). The third step localizes these edges. The strength image may consist of wide ridges, so "non-maximal suppression" is implemented to make edges one pixel wide, with output a binary image containing thinned edges. Finally, "hysteresis thresholding" is undertaken, which takes into account the directionality of the ridge by guessing continuation paths. Values along a ridge that fluctuate are compiled into a continuous lineation, with output a binary image consisting of detected edges (Figure 4.44c) (Trucco and Verri 1998). It is these edges that are used to create a final 3-D volume (Figure 4.44d), which can then be rotated and

examined at any angle (e.g., Figure 4.44e). This MATLAB 3-D model clearly shows the presence of three pit-room structures and another feature or structure in the southwest corner of the grid, which could possibly be another pit-room or series of connected rooms. The 3-D rendering informs us that it extends through numerous slices and changes shape and form with depth, as do several of the other anomalies.

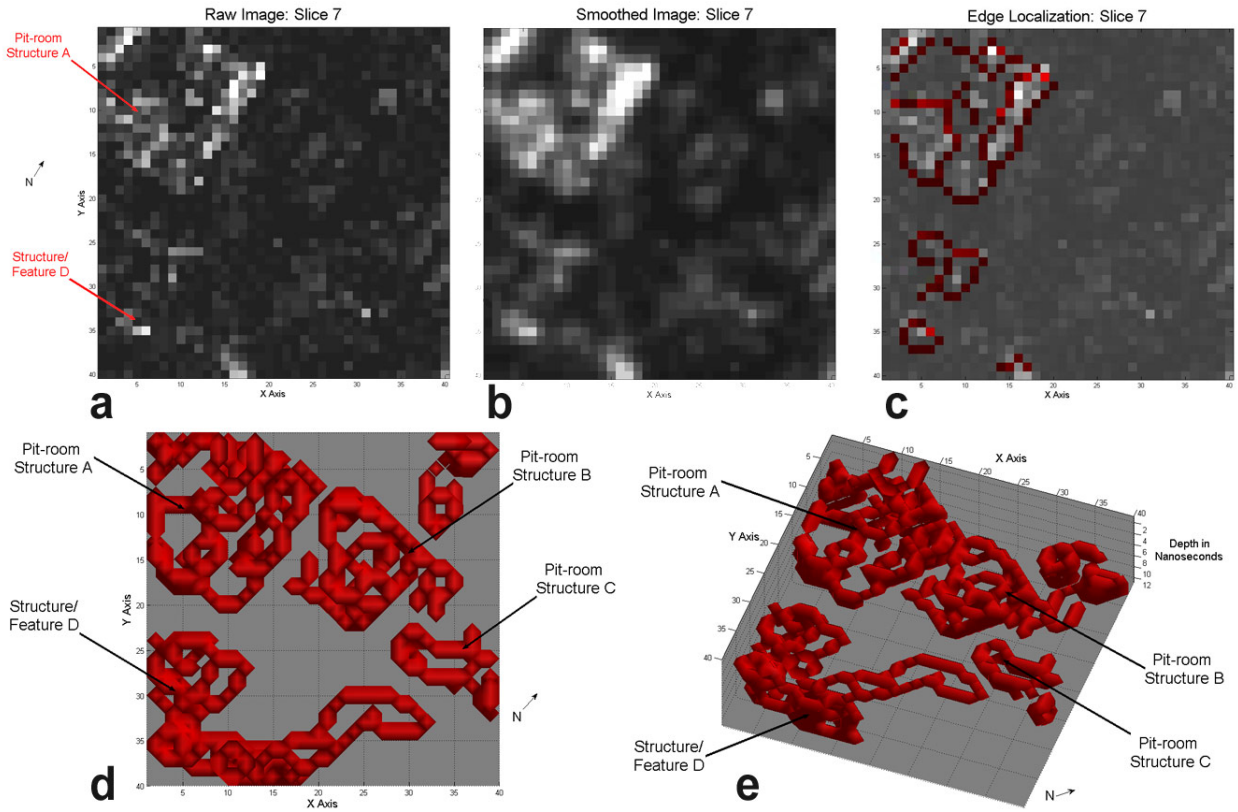


Figure 4.44. *3-D visualization with MATLAB 7.0 software using the Canny algorithm. a) Raw GPR Slice 7. b) Slice 7 after smoothing. c) Slice 7 after Sobel filtering and ridge detection. d-e) The 3-D data after incorporation of all 13 slices.*

4.6.3. Geophysical Results at Silver Bluff Plantation

Kenneth L. Kvamme, University of Arkansas

Geophysical data were collected at Silver Bluff by Archaeo-Physics, LLC, in November 2002. Initial guidance was given by Kenneth Kvamme, University of Arkansas, and Michael Hargrave, Construction Engineers Research Laboratories, in establishing the data collection grid within the site clearing (Figure 3.5). With extensive excavations already conducted at the site by Herron and Moon (2005), much was known about the character, depth *and orientation* of the site's historic archaeological features. Characteristically for historic-period sites, the many walls, post alignments, and other lineations that had already been excavated were aligned on north-south and east-west axes. It was therefore very important to rotate the geophysical survey grid at an angle relative to these features in order that survey transects would cross them at an angle. In this way, any survey defects would not align with archaeological features and processing

algorithms designed to eliminate them would not be at risk of removing anomalies of possible archaeological importance (see Section 4.4). Archaeo-Physics LLC utilized three geophysical methods in their survey of Silver Bluff (Maki 2003), with instruments and sampling densities described in Table 4.3.

4.6.3.1. Magnetic gradiometry

The magnetic gradiometry survey proved fruitless for the most part, as a result of the numerous cut steel-wire pin flags that littered the site, as discussed in Section 3.3.4.3. Long pin flag wires exhibit strong dipolar anomalies (paired black-white anomalies; yellow arrow, Figure 4.45a), while shorter metal fragments appear as numerous small monopolar point anomalies in the data. Of course, as a historic period site, many of these anomalies could represent significant iron historic artifacts. A rather large number of permanent archaeological datum, composed of steel rebar, was also dispersed through the site. These may be seen as larger anomalies 3-5 m in diameter, appearing mostly white (red arrow, Figure 4.45a).

Table 4.3. *Geophysical methods, instruments, sampling densities, and principal anomaly types at Silver Bluff.*

Geophysical Method	Instrument	Original resolu-tion (m)	Prospec-tion depth (m)	Unit of measurement	Principal indicated anomalies
Magnetic gradiometry	Geoscan Research FM-36	.125 x .5	1.5	nT/m	Recent & historic iron artifacts, burned areas
Electrical resistivity	Geoscan Research RM-15 twin-probe array	.5 x 1	.5	ohm/m	Linear features (lanes?)-negative; linear features (berms adjacent to lanes?)-positive
Ground-penetrating radar	Sensors & Software pulseEKKO 1000, 450 MHz antenna	.025 x .5	1 (100 ns time window)	amplitude (16-bit scale)	Walls, post alignments, ditches, builder's trenches

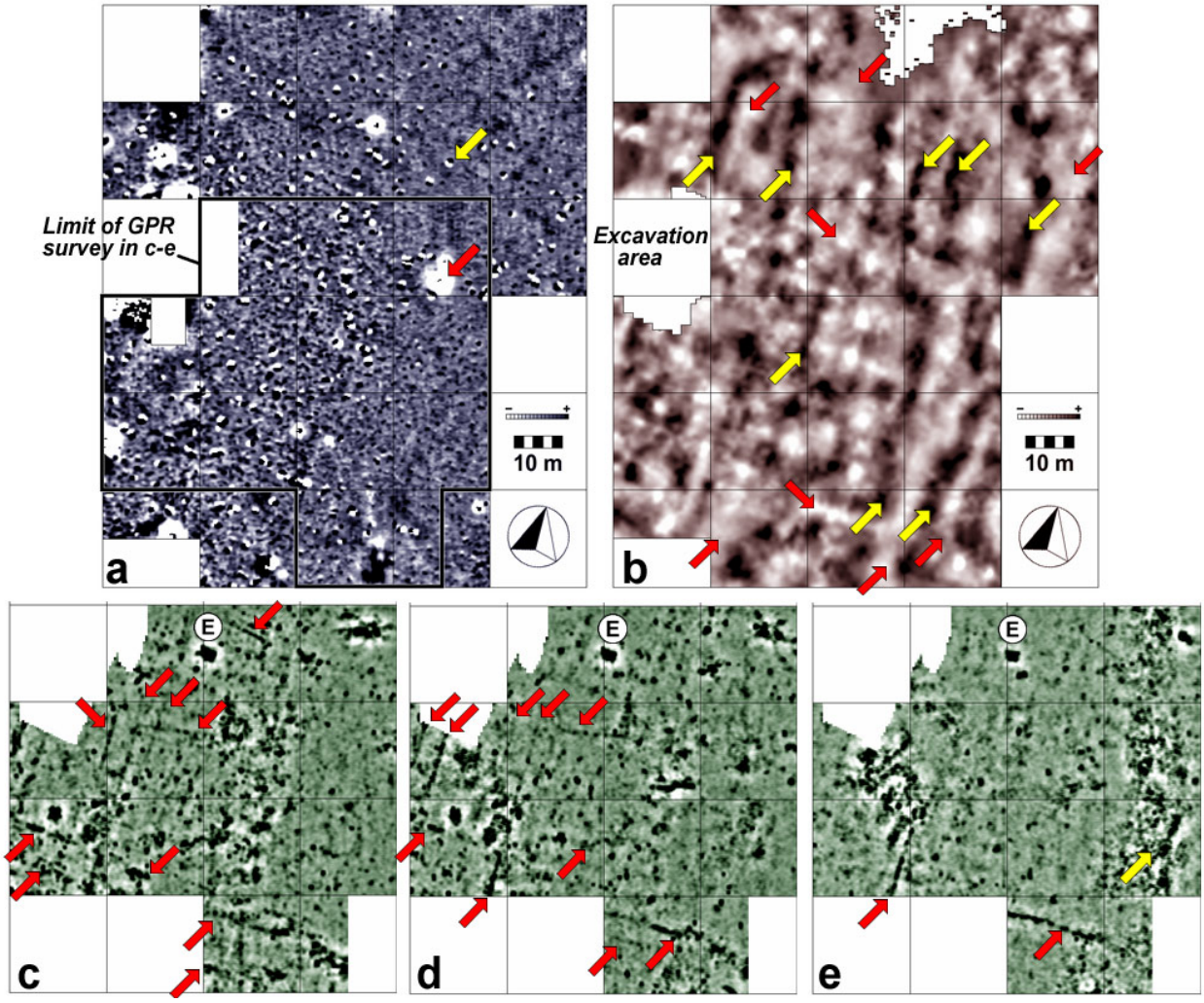


Figure 4.45. *Geophysical survey results at Silver Bluff Plantation. a) Magnetic gradiometry. b) Electrical resistivity. c) GPR depth-slice 40-60 cmbs. d) GPR depth-slice 60-80 cmbs. c) GPR depth-slice 80-100 cmbs. The label "E" signifies an open excavation pit. Arrows point to features discussed in text.*

4.6.3.2. Electrical resistivity

The electrical resistivity survey reveals a host of positive (yellow arrows) and negative (red arrows) lineations aligned on north-south and east-west axes (Figure 4.45b). Consequently, they parallel the many historic archaeological lineations discovered by excavation (Herron and Moon 2005) and also revealed by GPR survey (Figure 4.45c-e). Although these lineations are somewhat vague they appear to be cultural in origin—their orientation is correct and straight lines and apparent rectangular features do not appear in large numbers by chance. Their interpretation is a challenge, however, and it is uncertain whether it is the positive or negative anomalies that are culturally significant (or perhaps both might be). Stepping back, the data appear much like a grid of lanes with rectangular "blocks" of development between them. This suggests an interpretation that negative linear anomalies represent a series of lanes with the positive linear anomalies that parallel them representing a series of berms adjacent to the lanes. Small areas of negative anomalies may simply signify open ground between mounded soil. This view fits the

theory of resistivity where mounded earth is generally more resistant than depressions such as lanes. It also fits the historic evidence of significant development in this area with the presence of numerous structures already discovered and the likelihood unknown others in the region.

4.6.3.3. Ground-penetrating radar

The GPR data, offered in a series of depth-slices ranging from 40-100 cm in depth (Figure 4.45 c-e), support the previous resistivity-based model. Many lineations cross the site in north-south and east-west directions, indicating a substantial region of cultural constructions (arrows). Some of the GPR anomalies parallel positive resistivity anomalies, although a few coincide with negative resistivity anomalies (e.g., the linear anomaly by the lower right arrow in Figure 4.45e). Most, however, are independent, and offer much finer detail. The linear GPR anomalies appear to clearly represent wall of various thickness, at changing depths, and perhaps of different materials. Some may represent builder's trenches for walls or palisades. Other could represent closely spaced postholes. Broader GPR anomalies may signify prepared floors of brick, stone, or perhaps packed earth. The deepest GPR slice indicates the possibility of deep lineations to perhaps a meter below surface, perhaps indicating a deeper builder's trench or other similar feature. The broad zone of GPR disturbance in the deepest slice (yellow arrow, Figure 4.45e) appears to closely parallel a broad, negative resistivity anomaly (Figure 4.45b). It can only be hypothesized that a more robust, and deeper, lane may be indicated.

4.6.4. Geophysical Results at Kasita Town

Eileen G. Ernenwein, University of Arkansas

Geophysical surveys were performed at Kasita Town by Archaeo-Physics, LLC, in December 2002. Initial guidance in placing the survey area was given by Thomas Foster, of Panamerican Consultants, an expert on the site. Kenneth Kvamme, University of Arkansas, and Michael Hargrave, Construction Engineers Research Laboratories, conducted preliminary magnetic gradiometry surveys to site the location of the data collection area within a zone of apparent cultural anomalies (Figure 3.6). Magnetic gradiometry and electrical resistivity data were collected over a 1.2 ha area, and ground-penetrating radar with a .5 ha space inside that area (see Figure 3.6 where both survey areas are indicated). Instrumentation, sampling densities, and principal indicated anomaly types are given in Table 4.4.

American Revolution veteran Caleb Swan described Creek towns as extremely ephemeral, with houses built to last one to two years, after which new ones were built in different locations resulting in the shifting of entire villages along river ways (see Section 3.3.4.4). He also noted that areas where houses once stood only four years before were barely recognizable (Swan 1855: 692-693). This implies the archaeological record for these sites should consist of multiple post hole patterns superimposed on top of and adjacent to each other, with nearby clay extraction pits for mining raw materials for daub. This presents a challenge to the detection of Creek Indian structures by geophysical survey. Not only might posthole patterns be cluttered and incoherent, but postholes are also very difficult to detect geophysically. Furthermore, even if postholes are detected, they are very difficult to distinguish from background clutter produced by small objects

in the ground, rodent disturbances, and instrument noise. Only when enough posthole-size anomalies are found in a closely spaced geometric pattern can they be reliably identified as potentially cultural.

Table 4.4. *Geophysical methods, instruments, sampling densities, and principal anomaly types at Kasita Town.*

Geophysical Method	Instrument	Original resolu-tion (m)	Prospec-tion depth (m)	Unit of measure-ment	Principal indicated anomalies
Magnetic gradiometry	Geoscan Research FM-36	.125 x .5	1.5	nT/m	Recent & historic iron artifacts, hearths & burned areas, possible house edges, robust ditches /paleo-channels?
Electrical resistivity	Geoscan Research RM-15 twin-probe array	.5 x 1	.5	ohm/m	possible house edges, clay extraction pits, robust ditches /paleo-channels and ditch berms?
Ground-penetrating radar	Sensors & Software pulseEKKO 1000, 450 MHz antenna	.025 x .5	1 (100 ns time window)	amplitude (16-bit scale)	Post holes, recent & historic metal artifacts, robust ditches/paleo-channels?

4.6.4.1. Magnetic gradiometry

Potential features detectable at Kasita Town include burned houses, other structural remains, and large pits where topsoil has been removed and not replaced such as might occur in hypothesized clay extraction pits or burials. In addition, many of the artifacts associated with this historic period site could be made of ferrous metal, and so will yield strong dipolar anomalies. Some potential problems for magnetometry at the site include the likely presence of metal debris from airfield use, which introduces additional dipolar anomalies, and the possibility that the land has been plowed or graded, causing soil disturbance. The magnetic gradiometry results show many possible cultural anomalies along with clear disturbances from modern military activity (Figure 4.46a). Cultural features may be indicated by several linear and small circular positive anomalies, some of which form a house-like pattern (arrows, Figure 4.46a). The very robust, sinuous anomalies could possibly represent fortification ditches associated with Kasita Town, but since no other Creek villages are known to be fortified they more likely represent paleochannels. The linear anomaly composed of a string of dipoles crossing the center of the image from south to north (Figure 4.46a) is a ferrous metal pipe that crosses the site.

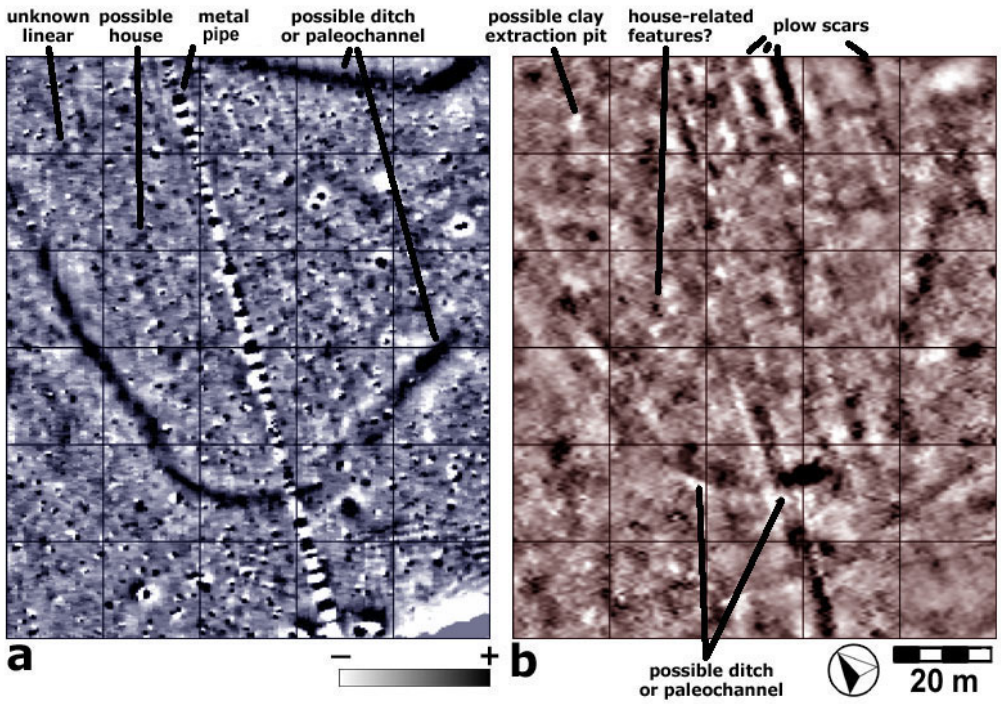


Figure 4.46. *Geophysical data collected at Kasita Town (a) magnetic gradiometry, and (b) electrical resistivity. Several initial anomaly interpretations are indicated.*

4.6.4.2. Electrical resistivity

If Kasita Town contains large clay extraction pits for building materials, and remains of house walls and foundations constructed from this clay, then both could possibly be detected by electrical resistivity. Deep plowing or grading of the area, as seems evident in the data (Figure 4.46b), presents a challenge to the detection of cultural features, particularly if the site itself is not deeply buried. The possible plow scars near the top of the image are strongly indicated, suggesting that they could run very deep. In addition, more subtle scars occur across the survey area at a right angle, possibly from earlier plowing or airport construction events (Figure 4.46b; some of these patterns were removed during data processing). The possible ditch or paleochannel that was so robustly indicated in magnetometry is also detected with resistivity, but to a lesser degree. When superimposed it becomes clear that the high magnetism of this features is directly associated with low resistivity, supporting the hypothesis of a filled ditch or channel. High resistivity anomalies parallel it to the north and most likely represent a berm. The trench that holds the pipe evident in the magnetometry data also is visible as a linear anomaly of high resistivity (Figure 4.46b). Possible cultural features in these data include negative anomalies that could possibly represent clay extraction pits plus a grouping of linear and small circular anomalies that could be related to the suspected house visible in the magnetometry data (Figure 4.46b).

4.6.4.3. Ground-penetrating radar

Of the three geophysical methods employed at Kasita Town, GPR should have the greatest potential for detecting postholes due to its higher sampling density (.2 m in the transect direction). Unfortunately, no clear patterns or alignments of small anomalies

suggestive of postholes can be seen in the data (Figure 4.47). Nevertheless, the many hundreds of small point anomalies distributed throughout the GPR data form possible candidates for postholes (Figure 4.47), although the majority are probably due to "noise" resulting from small objects (including metallic debris so evident in the magnetic data), rodent holes, or voids beneath the surface. The depth information provided by GPR shows that the ferrous metal pipe, which illustrates a pronounced anomaly, lies beneath the ground at about .6-.8 m depth (Figure 4.47d). In addition, the ditch or paleochannel feature is found in only the deepest slices, beginning at about 1 m below the surface and indicated most robustly in the deepest depth-slice (Figure 4.47f-h). Since clear geometric patterns are lacking in the majority of anomalies, further interpretation is limited and additional information can only be gained through excavation.

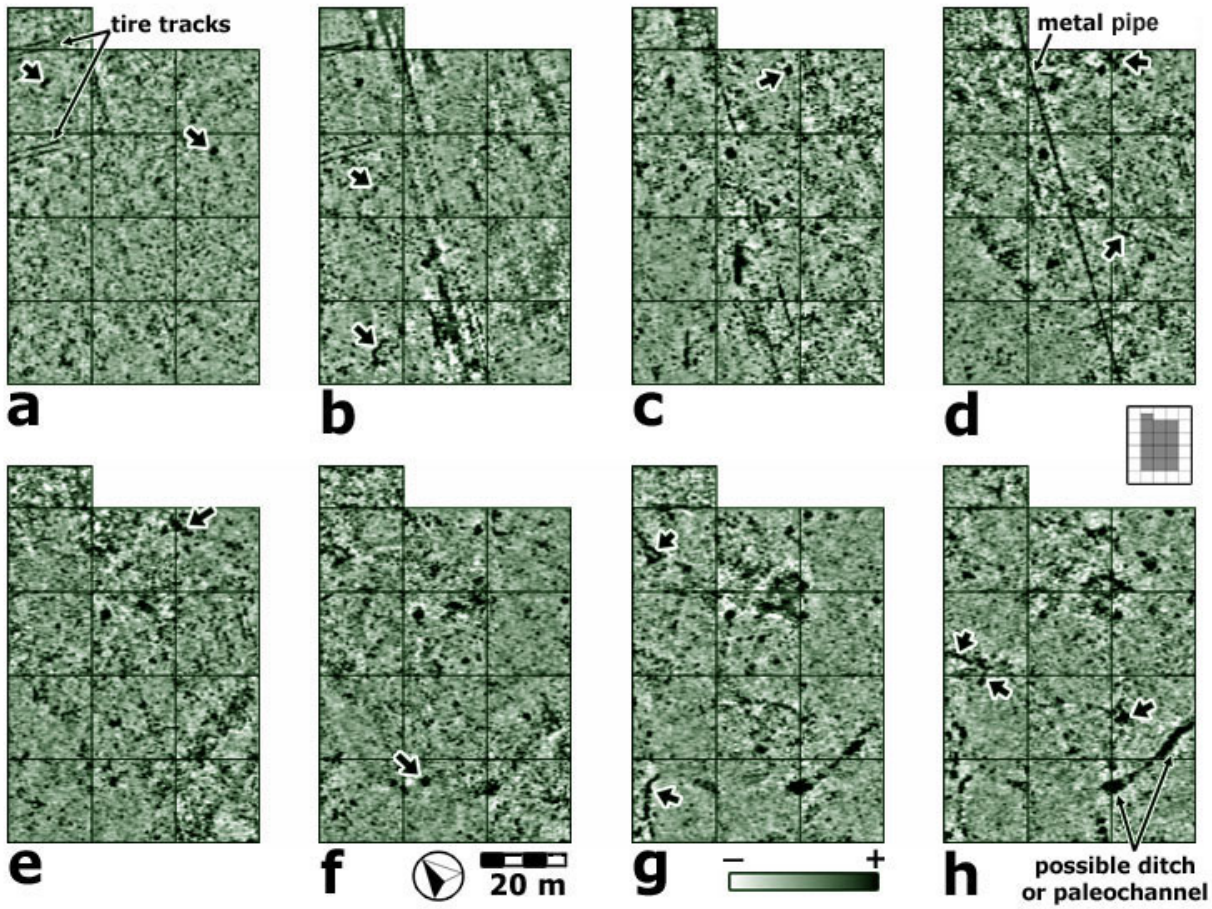


Figure 4.47. *Kasita Town ground-penetrating radar depth slices. Depth slices of .2 m increments beginning with (a) 0-.2 m and ending with (h) 1.4-1.6 m. Non-archaeological anomalies are labeled, and some of the many possible cultural features are designated with short arrows.*

4.7. AIR AND SPACE REMOTE SENSING THEORY & HISTORY

Kenneth L. Kvamme, University of Arkansas

Remote sensing from the air or space utilizes electromagnetic radiation reflected or emitted from the Earth's surfaces, offering a cost-efficient means for archaeological reconnaissance over wide areas (Bewley 2000:3). Aerial remote sensing uses airborne platforms while space remote sensing orbiting satellites. Both employ similar sensors with a principal difference lying in grossly divergent distances from the Earth's surface, which, until recently, has profoundly affected the spatial detail of what could be resolved from space. Aerial photography is the oldest and most-used domain of archaeological remote sensing, but other sensing devices have been placed in the air in recent decades including passive multispectral and thermal sensors making aerial remote sensing truly multidimensional. This section overviews the method and theory of the air and space remote sensing methods used in this project and presents basic findings.

4.7.1. Theory

The sun, source of most electromagnetic radiation (EMR) impinging on the earth, gives a spectrum ranging from biologically lethal gamma rays (short wavelength, high frequency) to passive radio waves (long wavelength, low frequency). For convenience, the spectrum is divided into regions possessing similar characteristics, "spectral bands." In order of increasing wavelength, the bands used in remote sensing include ultraviolet (UV), visible, infrared (IR), and microwave. Although EMR travels through empty space unimpeded, the earth's atmosphere allows certain wavelengths to pass freely (atmospheric windows or transmission bands) while others are blocked (atmospheric blinds or absorption bands), including short wavelength gamma, x-ray, and most UV radiation. The visible band (.4-.7 μm) is the region most commonly used with white light composed of three segments representing the additive primary colors blue (.4-.5 μm), green (.5-.6 μm), and red (.6-.7 μm). The more complex IR band is commonly divided into near (NIR: .7-1.5 μm), middle (MIR: 1.5-5.6 μm), and far (FIR: 5.6-1000 μm) regions, but also addressed as reflected IR versus emitted or thermal IR. On average, slightly more than half of the EMR hitting the earth reaches the surface, the rest being reflected or absorbed by the atmosphere. A small fraction of this energy is reflected from the surface as UV, visible, NIR, or MIR radiation, detectable from air or spacecraft. The remainder, absorbed by various surfaces, is transformed into low-temperature heat (felt as surface warming) that is reradiated continuously back into the atmosphere as long-wave thermal radiation (about 3-1000 μm, in the MIR and FIR ranges; Avery & Berlin 1992:3-14; Sabins 1997:2-6).

Any surface reflects or emits EMR over a range of wavelengths in a unique manner depending on its physical composition and state. The average amount of incident radiation reflected by a surface in some wavelength interval is its spectral reflectance or *albedo*, while emitted radiation yields information about temperature properties. This distinctive spectral response pattern, diagnostic of particular surfaces and their states, forms a *spectral signature*, allowing surfaces to be recognized from their spectrum. Water tends to reflect blue wavelengths and almost no NIR; most healthy vegetation reflects some green light but is enormously reflective in NIR; and many distinctive properties of soils and geological deposits are discriminated in NIR and MIR bands

(Avery & Berlin 1992:14-15). Sensor type determines the regions of the spectrum utilized. Some sensors provide photographs, pictorial representations recorded directly onto photographic film; others generate an "image," a more general term referring to all pictorial representations of remotely sensed data. All photographs are images, but not all images are photographs.

4.7.1.1. Thermal infrared

The thermal imaging of archaeological sites is a rarely explored topic, with little published literature in the Americas (see, however, Berlin et al. 1977 for an early case study). Significant research in this topic has occurred in France, however, particularly in the 1970s, where the enormous potential of this method for imaging buried architecture and other archaeological components was demonstrated (see Dabas and Tabbagh 2000 for a summary). Radiation is emitted by all objects with spectral properties and magnitudes determined by temperature and the material's *emissivity*, its efficiency as an absorber and emitter. As temperature decreases, emitted energy shifts to longer wavelengths: the sun's surface temperature of about 6000° C emits peak radiation at a wavelength of .48 μm, seen as light, while the earth's ambient temperature of about 27° C yields a maximum emissivity at 9.7 μm, felt as heat. The latter wavelength conveniently falls within one of only two "windows" within which thermal infrared radiation is not blocked and can be remotely sensed in the lower atmosphere (Avery and Berlin 1992:120). The thermal emittance of an object that defines its radiant temperature can be measured by a non-contact thermal infrared radiometer, with resolutions in the range of .01-.25° C. Such a device, on an airborne platform, is essential to applied work because surface temperatures vary so rapidly through the day that slow-to-acquire ground measurements of temperature are unsuitable if one wants to measure real spatial changes instead of temporal ones.

Thermal processes in archaeological sites are complex. They are affected by *thermal conductivity*, the rate at which heat flows through a material, *thermal capacity*, the ability of a material to store heat, *thermal diffusivity*, the rate of heat transfer within a substance, and *thermal inertia*, the tendency of a material to resist temperature change. In general, dense materials like rock possess greater thermal inertia than porous, unconsolidated materials. Porous materials (e.g., dry sand) possess low thermal inertia and readily change temperature. Rock and compacted deposits, on the other hand, resist temperature change. Thermal energy from the sun, the atmosphere, or conductance from surrounding deposits causes the ground to heat. The amount of thermal energy in the ground depends on the balance between these forms of gain, and of loss by re-radiation, evaporative cooling, air convection, or diffusivity. The diurnal cycle of warming in the day and cooling at night principally affects only a near-surface "skin" layer, perhaps to a depth of 15 cm. Thermal variation in this zone is affected by aspect, because slopes facing the sun become warmer than shadowed slopes (an effect that can last for hours), by evaporative cooling effects (moist ground tends to be cooler than dry ground), wind effects, and variations in rates of evapotranspiration caused by changes in plant cover. Thermal energy in deeper deposits, principally transferred by conduction, tends to be affected by longer-term cycles, primarily weather changes of several days where temperature is consistently rising or falling. The depth limit of thermal prospecting is perhaps a meter, dependent on the duration and intensity of a thermal cycle, the volume

of the feature to be detected, and its thermal inertia contrast against the surrounding matrix (Dabas and Tabbagh 2001). The combined effects of these processes produce a complex outcome, but clear spatial structure within thermal infrared imagery facilitates interpretation.

4.7.2. Archaeological Aerial Photography

4.7.2.1. History

Aerial photography was pioneered from balloons in the 19th century, with the earliest aerial photograph of an archaeological site, Stonehenge, taken in 1906. World War I (1914-18) saw vast improvements with a move to maneuverable aircraft and trained personnel, of whom several focused on archaeological sites during the war. Outstanding aerial photographic surveys of Roman occupations were carried out in Syria during the 1920s-30s and in North Africa during the 1940s. O.G. S. Crawford, an observer in the British Royal Flying Corps, was primarily responsible for the early development and systematization of aerial archaeology. While sites and features in arid lands were revealed largely by shadows cast by crumbling ruins, Crawford recognized prehistoric structures in air photos taken in England by "crop-marks," a means of detection so successful that he documented more archaeological sites in a single year than had been located by pedestrian surveys in the previous century (Bewley 2000:3-4; Wilson 2000:16-20). Aerial photography is now a critically important tool for archaeology, particularly in the United Kingdom, France, and Germany. Bewley (2000:4) suggests it caught on in Britain through the nature of the landscape and its archaeology, the ability of the soils to produce excellent crop-marks, the availability of aircraft, relatively free airspace, and a military establishment that encouraged archaeological discoveries. In these countries aerial photography is closely integrated with other techniques as a branch of archaeological survey; there are wide areas of coverage, cost effectiveness, and a landscape perspective in the approach to sites. Extensive archives and libraries of archaeological aerial photographs have holdings numbering in the millions; the British National Association of Aerial Photographic Libraries lists more than 360 libraries (Bewley 2000:8). European aerial archaeology is highly organized, with state-sponsored agencies, funding, professional organizations, and significant archives. The systematic coverage of landscapes remains the principal goal, with resurvey and revisits to monitor changes and new conditions (Whimster 1989). Wilson (2000) argues that aerial photography is responsible for the discovery of more archaeological sites than all other methods, pedestrian or remote sensing, combined.

In North America the potential of an aerial perspective was also seen very early, with photographs of Cahokia in 1921, then Lindberg's photographs of Southwestern pueblos and sites in the Yucatan peninsula in the late 1920s and early 1930s (Avery & Berlin 1992:226-227), and a Smithsonian Institution survey made over 700 aerial photographs in the Gila River Valley, Arizona, to document ancient Hohokam irrigation canals (Reeves 1936:103). Despite this, serious attention and organization of aerial archaeology has not been realized in the United States (with the possible exception of Chaco Canyon National Monument, New Mexico: Drager & Lyons 1985), where it is rarely incorporated into regional projects, research designs, or in state-sponsored inventory programs. Nevertheless, the few examples of North American work together with the great success of this method in Europe point to undeniable potential for mapping

hidden and buried characteristics of archaeological sites. This potential is realized through several mechanisms.

4.7.2.2. Archaeological air photo interpretation

The basis for assessing most forms of aerial or satellite imagery lies in concepts of pattern recognition of cultural forms based on their regular geometric shapes and systematic patterns (Section 4.5.3). There are four general classes of phenomena originally codified by Crawford and others that generate these patterns and permit identification of archaeological sites and their internal features (Scollar et al. 1990:33-74; Wilson 2000:38-87). Unlike other forms of aerial evidence, these phenomena are nearly unique to archaeological identifications.

Vegetation or crop mark sites are caused by the differential response of plants growing on the surface to what lies buried beneath. It is the principal reason why archaeological sites can be seen from the air. Each plant acts as a kind of subsurface sensor, with roots penetrating to a depth roughly equal to its height above ground and sensitive to particular soil characteristics in its immediate vicinity—moisture, material composition, nutrients. In an agricultural field of uniform crop type, literally millions of similar and uniformly placed plants are distributed over a broad area. When viewed from the air, changes in plant size and color can be readily detected that arise from two possible impacts of buried archaeological features on plants. *Positive vegetation marks* typically occur when plants grow over a richer, wetter soil or moisture trap that results in advanced growth, taller or more robust plants, color changes relative to other plants, or increased densities (Figure 4.48a). These conditions usually point to archaeological features like ditches, trenches, pits, or depressions of any kind, such as those associated with former dwellings or cellars. *Negative vegetation marks*, on the other hand, are caused when archaeological features near the surface—stone or brick walls, floors, plaster or tile concentrations, packed earth, pavements—retard plant growth by reducing nutrients, moisture, or other factors (Figure 4.48b). Typical effects are color differences, height variations, or density changes seen as yellowing, stunting, wilting, or delayed growth (Wilson 2000:67-87).

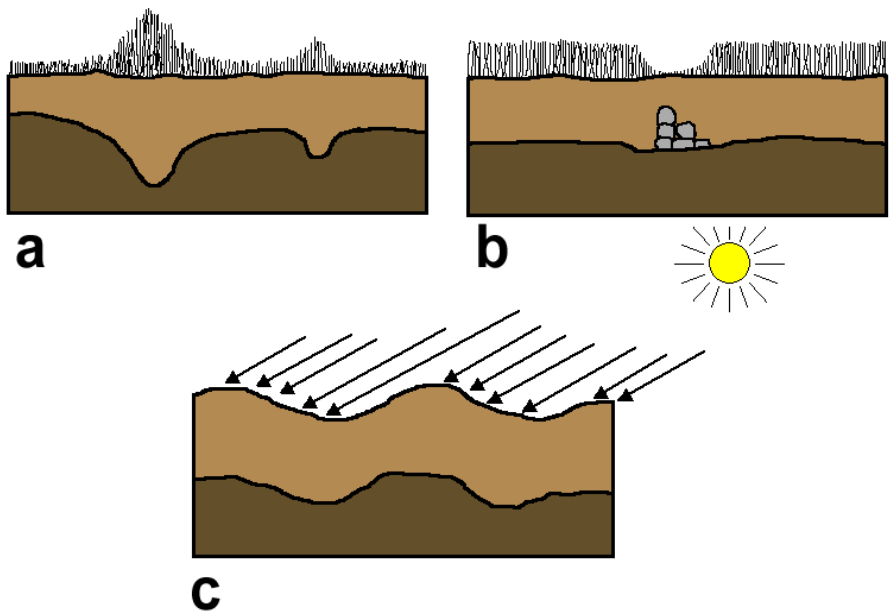


Figure 4.48. *Phenomena that cause features of archaeological sites to be visible from the air. a) Positive vegetation or crop marks caused by advanced growth over prehistoric ditches. b) Negative mark caused by retarded growth over stone architecture. c) Shadow marking caused by subtle changes in microtopography.*

<u>*Shadow mark sites*</u> indicate elevation changes in the surface caused by standing ruins, structures, or mounds, but more often by subtle rises over shallowly buried architectural features and depressions over ditches, pits, roads, or trails (Figure 4.48c). They become visible under low sunlight illumination angles (e.g., between 10-30°). Shadows on slopes facing away from the sun immediately adjacent to bright sunward-facing slopes exaggerate whatever landscape hints might be present, many of which may not be readily visible on the surface (Bewley 2000:5).

<u>*Soil mark sites*</u> arise from variations in soil color in open fields, particularly freshly plowed ones, caused by differences in mineralogy or moisture content. Where imported foreign materials (bricks, plasters, clays), introduced organics, baking of soils, or in-filling of ditches or depressions has occurred, color, texture, or soil lightness changes are possible. Changes in soil moisture can enhance these effects where fills within ditches or depressions might dry out more slowly, yielding an observable brightness or texture difference. Differential soil drying can also result from microrelief expressions where high points loose moisture more quickly than low (Scollar et al. 1990:37-48).

<u>*Snow and frost mark sites*</u> are revealed by two circumstances. Variations in thermal energy emitted by archaeological deposits may cause differential melting of frost or a light snow (see discussion of thermal infrared below). Microrelief also produces differential melting because snow or frost on slopes facing the sun disappears more quickly than on shadowed slopes. From the air, both can reveal archaeological patterns distributed through the landscape (Scollar et al. 1990:48-49; Wilson 2000:43-46).

Several examples of crop, vegetation, and shadow marking of archaeological sites are given in Figure 4.49 that demonstrate the enormous utility of an aerial vantage for understanding the extent, composition, and internal organization of past settlements and other constructions. Combined with the advantages of ground-based geophysical operations, it should be apparent that both domains of remote sensing can potentially contribute enormously to a more complete and richer understanding of the archaeological subsurface.

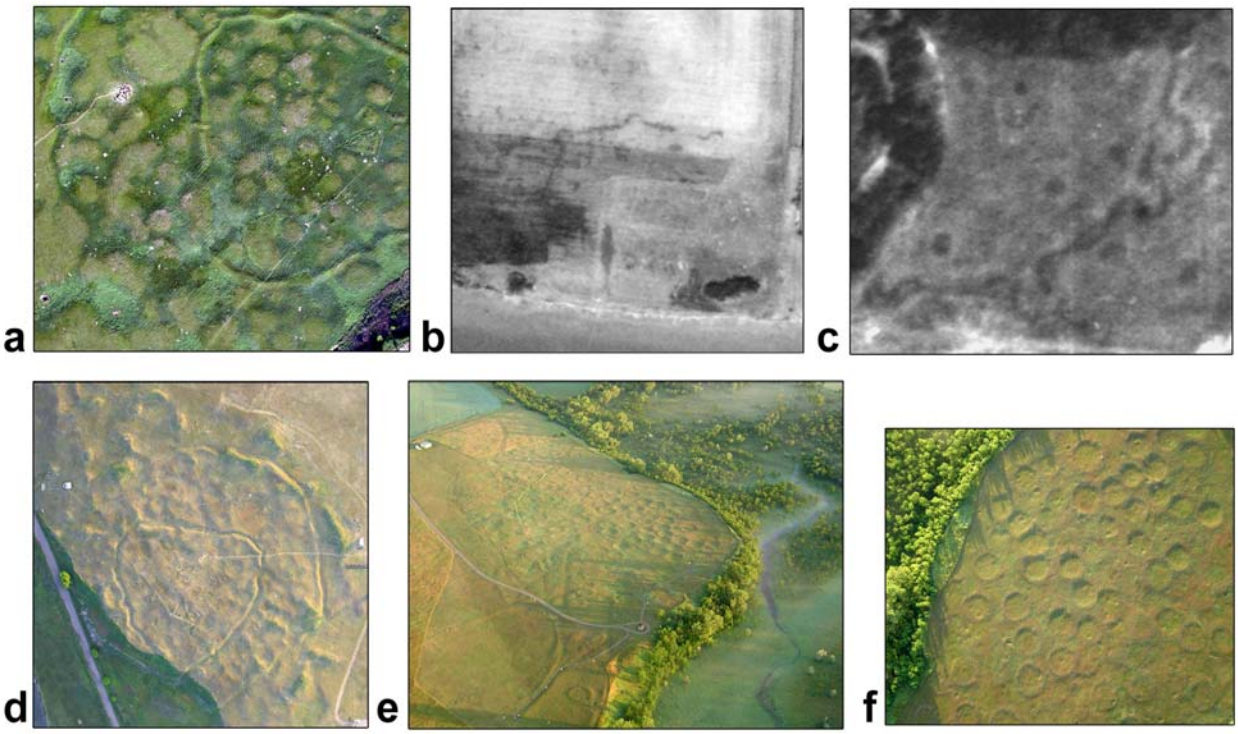


Figure 4.49. *Vegetation and shadow marks point to subsurface archaeological conditions. a) Vegetation marks distinguish between houses, village fill, and middens at 15th-18th century Double Ditch Village, North Dakota. b) Fortification ditch, bastions, and several houses indicated by positive vegetation marks in a 1968 soil survey photo of 14th century Whistling Elk village, South Dakota. c) Fortification ditch, bastions, and several houses indicated by positive vegetation marks in a 1938 soil survey photo of 13th century Menoken Village, North Dakota. d) Shadow marking in a sunrise view of Double Ditch. e) Shadow marking shows houses, trails, defenses, and corrals at the 19th century Mandan-Arikara village at Fort Clark, North Dakota. f) Close-up of houses and collapsed subterranean storage pits at Fort Clark.*

Critical to the foregoing recognition domains of aerial (or space) remote sensing are clear weather, good visibility, and the timing of flights. Sunrise or sunset flights during clear sunny days are necessary to record shadow sites; crop marks must be timed to crop development, itself dependent on plant type, growth cycle, drought, or rainfall abundance; soil marks are best seen after a fresh plowing when soil moisture remains high; a light snow or frost is required to reveal snow and frost markings—but these must be caught before the sun has melted them! Repeat visits at multiple times of the year can

reveal different features, as can photographs of the same site from varied angles and lighting conditions (Wilson 2000:80-83).

4.7.3. Aerial and Space Imaging Activities in the SERDP Project

With obvious benefits of an air or space vantage and the excellent ground-based geophysical results obtained at several of the project sites (described in Section 4.6) the project team was excited by the potential of combining high quality remote sensing from both domains to form effective fusions of the data. Unfortunately, as the project progressed it became obvious that a combination of the kinds of sites selected for the project, timing, and environmental conditions worked against the acquisition of high-quality imaging of archaeological features from the air or space. None of the selected project sites contained microtopographic or other relief changes that might allow shadow marking or snow and frost marking. Furthermore, none occurred in cultivated fields that might allow visualization of crop marks or soil marking variations after a fresh plowing. That left vegetation marking, plus thermal variations (discussed below), as the sole possibilities that might allow detection of archaeological features from the air or space. Yet, it must be realized that in Europe there is a corpus of aerial specialists (many state supported) that constantly monitor fields holding important archaeological resources. They can time flights within days (or even hours) of optimal conditions and they conduct regular aerial reconnaissance flights (Wilson 2000). In our project we did not have that luxury. Owing to funding limitations we could acquire only single satellite scenes from databases of extant scenes and time single flights for aerial data collection. Consequently, although we attempted to estimate best times for flights and scenes, the aerial and satellite data utilized by the project are necessarily sub-optimal and do not fully represent the enormous potential of an aerial vantage. It is a major concern that the lack of high quality air or space results in this project will be interpreted as indicating that aerial methods are not worth pursuing in future archaeological remote sensing projects. Nothing is farther from the truth. Aerial methods have been shown to yield great results at many kinds of archaeological sites, and frequently they do so consistently. If they do not produce results, then it may only be a matter of poor timing.

Our primary consultant for the air and space phase of the project was Dr. Thomas Sever of the Marshall Space Flight Center of NASA. That team acquired a host of aerial and space data designed to point to subsurface conditions. The University of Arkansas geophysical team also had limited access to an aerial platform and was able to conduct limited flights at two of the project sites to augment NASA findings. The following sections describe these efforts and findings.

4.8. AERIAL & SPACE REMOTE SENSING BY UNIVERSITY OF ARKANSAS

The University of Arkansas team acquired limited aerial data at two of the project sites: Army City and Pueblo Escondido. This section describes these aerial activities and results.

4.8.1. Powered Parachute & Instrumentation

Kenneth L. Kvamme, University of Arkansas

Aerial flights at Army City made use of a powered parachute, in April, 2004. This device is a slow and low-flying ultralight aircraft piloted by Dr. Tommy Ike Hailey of the Cultural Resource Office, Northwestern State University, Louisiana. It is a two-seat aircraft that allows the pilot to operate from one seat, and permits a second person to freely operate cameras (Figure 4.50). Its top speed is about 30 mph, but that is generally reduced when flying into a headwind (the usual practice) such that the aircraft may be nearly stationary in flight. This greatly reduces image blurring from platform motion, allowing imagery of extraordinary detail (see comments about motion problems by NASA team in Section 4.7.5). The powered parachute can fly extremely low, with altitudes typically varying between 100-500 meters above the ground, allowing spatial resolutions of only a few centimeters (Hailey 2005).

The powered parachute was employed to acquire digital color imagery (still frames and video), as well as thermal infrared video. A Minolta Dimage A2 eight megapixel camera was utilized to capture most of the still color imagery, while a Sony DCR-TRV900 was used to capture color digital video. A Raytheon Palm-IR250, with sensitivity to about 0.1^{o}C, was used for thermal infrared data collection. This instrument captures data onto a relatively small CCD array (320 x 240) that is resampled to produce images of size 640 x 480. Eight-bit data from the 10 micron band allow warmest temperatures to be assigned a value of 255, with coolest temperatures quantified as zero, representing white to black, respectively, in the imagery. A digital tape recorder connected to the Palm IR-250 allows information to be recorded continuously at a rate of 30 frames/second during aerial flights. Still frames can later be extracted from the digital videotape and assembled as a mosaic to create larger composite imagery showing features of interest.

Figure 4.50. *Two-seat powered parachute piloted by Dr. Tommy Ike Hailey at Army City. The individual in the second seat focuses on data collection.*

4.8.2. Orthorectification

Kenneth L. Kvamme, University of Arkansas

The original coordinate grid at Army City was established by total station. This coordinate grid covered the entire field for purposes of the field-wide electrical resistivity survey (Hargrave et al. 2002), with a wooden stake placed every 20 m. An intensive effort was spent at the initiation of fieldwork in 2002 to relocate these datums. This was successful, and perhaps 75 percent of the badly decayed wooden stakes were found. Surveyor grade kinematic GPS units were employed to obtain real world coordinates for a suite of these datums defining the perimeter of our 100 x 160 m survey area. Permanent plastic datums (tent stakes) were sunk every 20 m around the site perimeter.

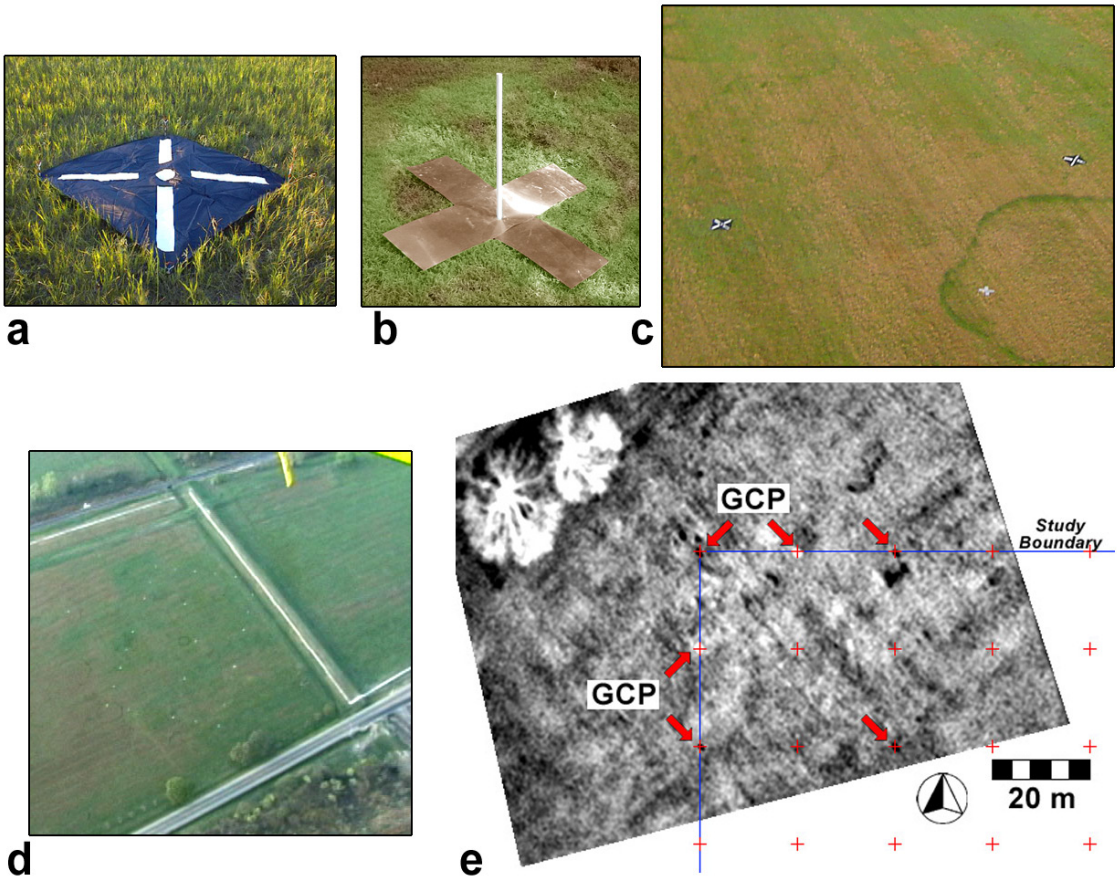


Figure 4.51. *GCP and image rectification and registration. a) Plastic GCP. b) Metal GCP. c) GCP from low altitude. d) GCP revealing the entire study area at Army City. e) Example of rectification and registration of thermal infrared scene onto coordinate grid (red crosses) at Army City.*

Nylon (Figure 4.51a) or metal (Figure 4.51b) crosses were placed over selected datums as ground control points (GCP) for purposes of image rectification and registration (Burrough and McDonnell 1998). The metal crosses were made of heavy-duty aluminum roof flashing and employed under the theory they would be more visible to thermal infrared imaging. They were, but so too were the nylon GCP. Both types were

placed alternately around the perimeter (Figure 4.51c) and in the interior of the study site (Figure 4.51d). Study area-wide mosaics were built using image rectification and registration techniques in Adobe Photoshop and GIS methods that "rubber-sheeted" slightly oblique imagery into a correct spatial projection in the site's local coordinate base. This process is illustrated in Figure 4.51e where a thermal infrared frame is "fit" to the local coordinate base.

4.8.3. The Potential of Aerial Remote Sensing at Army City

Kenneth L. Kvamme, University of Arkansas

It was emphasized above that if aerial remote sensing methods do not yield useful results that point to archaeological features that it may be only a matter of poor timing. Flights in other season or climate regimes (e.g., droughts or periods of heavy rain) could yield very different outcomes. This section focuses on this issue and demonstrates that exceptional aerial imaging of Army City's underlying structure and content would have been highly likely if aerial survey could have been undertaken at the proper time.

The summer of 2002 was a time of severe drought in much of the Great Plains, including northeastern Kansas. At the time of the July geophysical surveys, with temperatures consistently hovering near 40° C, the ground was extremely dry and parched, which made geophysical surveys difficult (see Section 4.6.1). An unforeseen consequence of the drought conditions, not fully appreciated at the time, was that the site's surface vegetation was extremely stressed, resulting in strong *negative vegetation markings* that correlated with many subsurface archaeological features of Army City indicated by the geophysics. In other words, one could discern the lines of walls, streets, and even individual rooms on the surface by merely walking about the site. This was realized as an important signal at the time, and a series of surface photographs were made from a borrowed two-meter-tall stepladder of several of these markings. It was believed at this time that air or space imagery to be subsequently acquired would capture these or similar markings. It was *not* fully realized or appreciated that strong vegetation marks were rare at this site—they probably occur only in drought years or periods of extreme dryness. More than a half-dozen subsequent visits and an equal number of air or spaces images confirm this view (see reviews of imagery below). It is unfortunate, then, that under the premise that wet spring conditions with new plant growth might yield *positive* vegetation marks and a "chlorophyll response," particularly in near-infrared bands, aerial imagery of Army City was acquired in April, 2004, and from the QuickBird satellite in a scene dating from April, 2001 (see below). At these times little anthropogenic patterning can be discerned in the imagery. As indicated by Wilson (2000) and others, vegetation markings expressed by archaeological sites have an idiosyncratic character dependent on unique site and climate conditions. At Army City it is now apparent that drought conditions *must* be sought to reveal the subsurface through vegetation markings. This turn of events has been a grave disappointment to the project, given the high quality of geophysical results. The following surface images demonstrate that aerial imaging could, indeed, be very productive at this site if properly timed with dry or drought conditions.

During the ground-based geophysical surveys of Army City, performed July 8-19, 2002, numerous surface features of potential interest to the interpretation and understanding of the site were observed. The field had been mowed immediately before

the survey, which facilitated the visibility of new plant growth that occurred after the brief and freak shower on the night of July 10. Some vegetation marks were made visible by new plant growth, but most were discernable as areas of stunted or wilted growth, or more often dead plants or no plants. An example of a likely cultural feature revealed by new plant growth is shown in Figure 4.52. It reveals a rectangular are measuring about 2 x 4 m, supporting a lush growth of new, green grass. It is likely caused by a moisture trap and a small region of better soil, probably within an individual room.

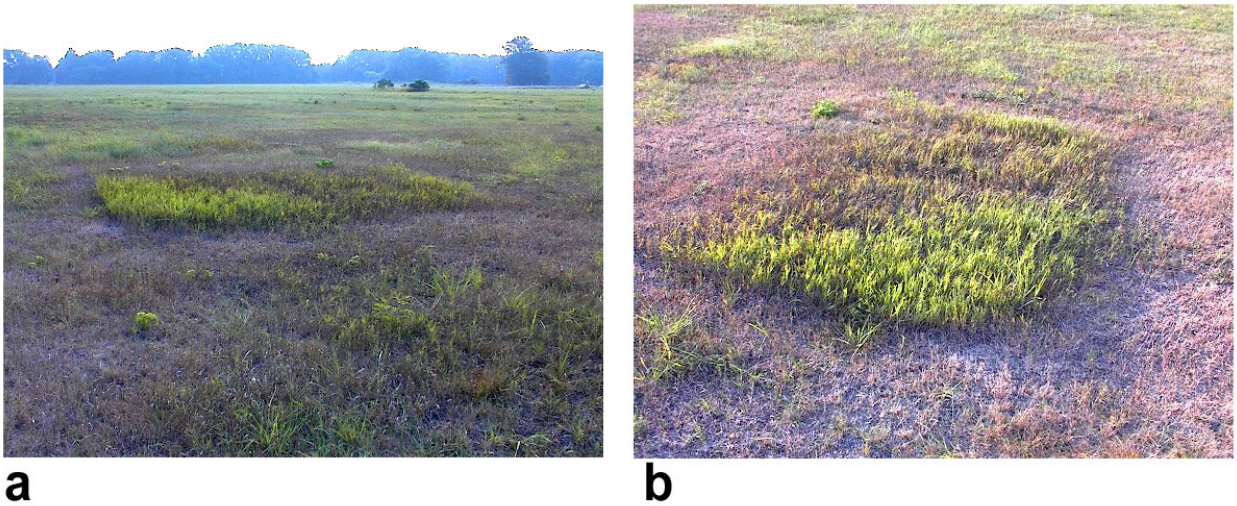


Figure 4.52. *Rectangular cultural feature of Army City, probably a room, indicated by positive vegetation marking in the form of new grass growth. a) View to east. b) View to south.*

Most cultural features at Army City were revealed by stunted or browned grasses. A good example is a portion of historic General Street, which is seen as a broad area of stunted grass (between arrows, Figure 4.53). This view is immediately south of the historic Hippodrome, and west of Washington Avenue, Army City's principal road (see Figure 4.37e).

Figure 4.53. *Stilted and browned grasses forming negative vegetation marks that indicate a segment of General Street, between arrows. This view looks east, immediately to the south of the Hippodrome and west of Washington Avenue.*

Remarkably, numerous "walls," some only 25 cm wide, were also apparent in the grassy surface of Army City. They were typically seen as linear features of browned grasses or were discerned as narrow areas lacking vegetation (Figure 4.54). Most of these walls correspond with the area of the historic Hippodrome (see Figure 3.2), and represent elements of this structure. Frequently, the patterning is clear enough such that individual rooms of this structure can be discerned. Subsequent excavations made in 2004 showed these walls to lie only 20 cm below the surface and that they were made of concrete (see Section 5.16).

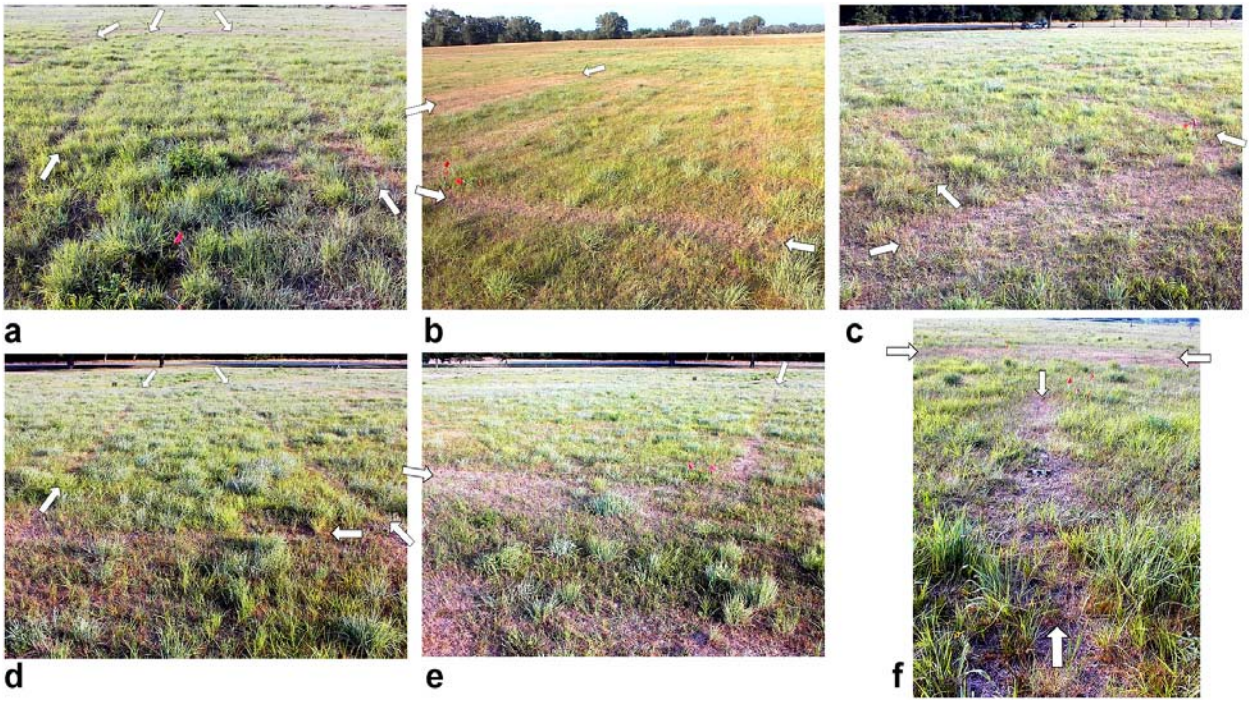


Figure 4.54. *Negative vegetation marks revealing linear architectural features representing walls at Army City. a-e) Views of browned and wilted grasses over the concrete walls of the Hippodrome, lying only 20 cm below surface. f) Hippodrome wall with 25 cm scale (center); Colonel Street may be seen between horizontal arrows.*

Even more impressive and convincing, however, were dozens of small, circular areas of dead or browned grasses, about a half-meter in diameter (Figure 4.55). They occur in linear arrangements and clearly represent building footings. Subsequent excavations made on 2004 indicated brick and cement footings lying only 20 cm below the surface (see Section 5.16), which evidently stressed plant life growing over them.

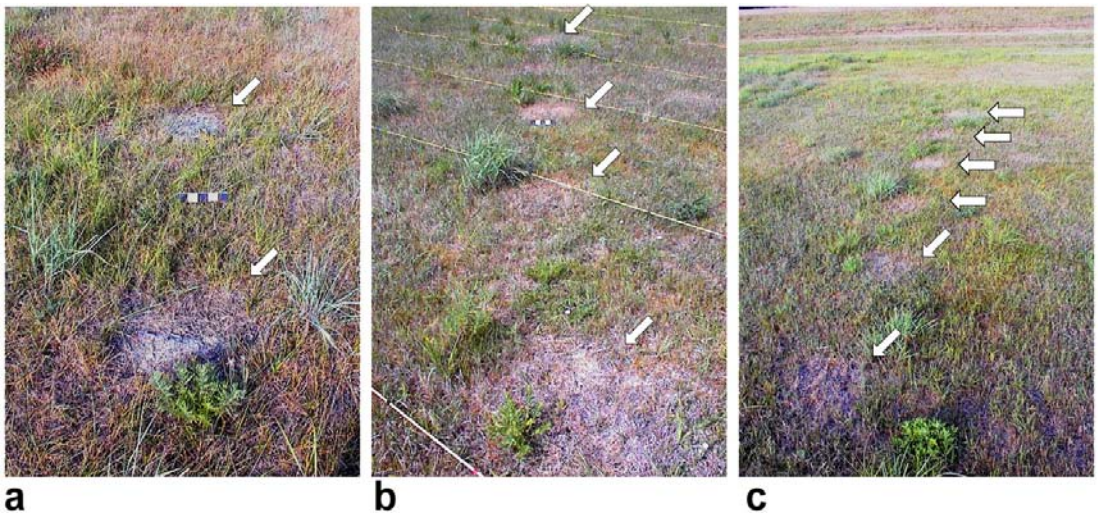


Figure 4.55. *Building footings revealed as negative vegetation markings at Army City. a) Two footings (arrows) with 25 cm scale. b) A linear arrangement of four footings (arrows) with geophysical survey guide ropes (each separated by 2 m) placed for survey. c) Six building footings (view to southwest, with levee in background).*

It is very clear from Figures 4.52-4.55 that had the site been viewed and imaged from the air at the time of the geophysical surveys an extraordinary aerial mapping of Army City would have been achieved. In fact, such a mapping was attempted of what could be seen from the limited standpoint of the surface. It is standard practice of the ArcheoImaging Lab to systematically inspect and map the ground surface during geophysical surveys because circumstances visible on the surface (e.g., badger dens, rocks, trees, bushes) frequently explain anomalies later observed in the geophysical data.

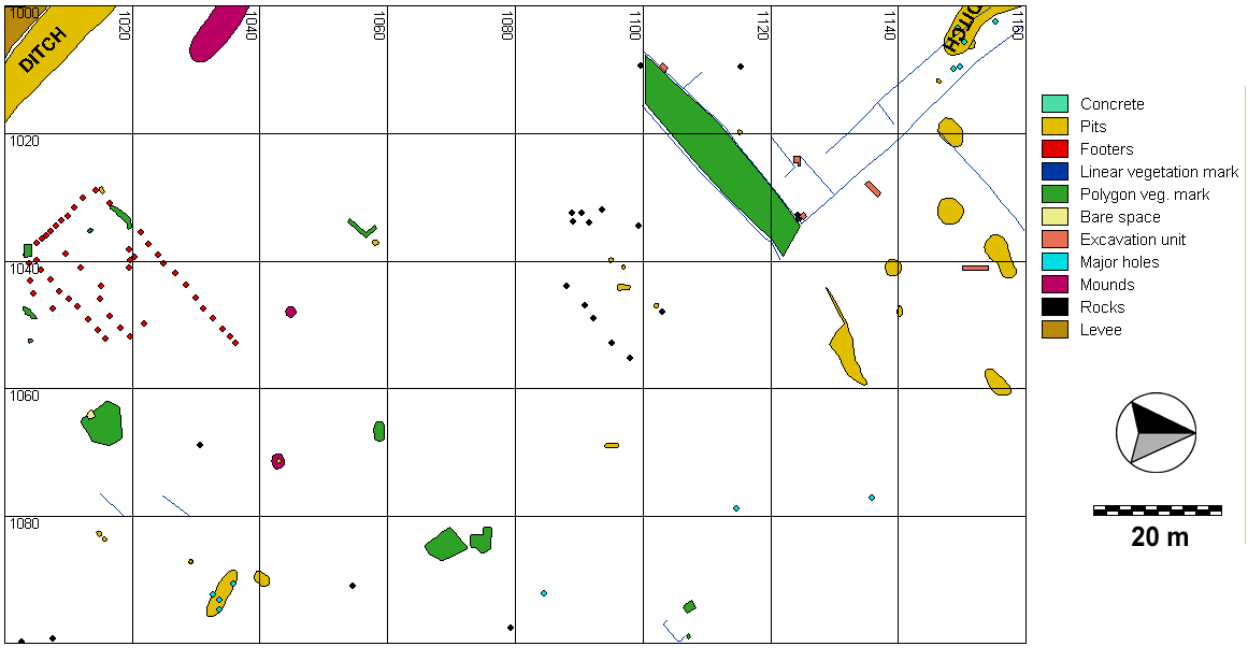


Figure 4.56. *Map of surface-visible features, many revealed by vegetation markings, made at the time of the geophysical surveys.*

Maps of each 20 x 20 m survey block were therefore made throughout the 100 x 160 m study region. They depict every observable nuance or impression visible on the surface. A compilation of these maps for the full study area is illustrated in Figure 4.56. It shows much of the layout of the Hippodrome, the outline of a structure supported by footings, rocks jutting from the surface of a rectangular structure, and many depressions in the vicinity of the former Orpheum Theater (see Figure 3.2).

4.8.4. Normal-light Aerial Photographs and Imagery from Army City

Kenneth L. Kvamme, University of Arkansas

A number of aerial scenes were procured from Army City, but in the collective set few archaeological features are indicated, if any. A National Aerial Photography Program (NAPP) aerial photo dating to 1991 shows a number of large trees near the study area (appearing black from extreme contrast manipulation). The area of the hippodrome (upper left of study area) and a number of subtle linear/rectilinear anomalies may be visible (Figure 4.57a). A NAPP photo from 1996 shows a similar but less crisp view. A lineation in the upper center probably represents Washington Avenue, a square anomaly to the left possibly indicates a buried structure (Figure 4.57b). A low resolution color image made at high altitude for soil surveys in 1999, obtained from the Fort Riley Directorate of Environment and Safety, shows the trees are probably down and replaced by brush (Figure 4.57c). Heavy disturbance obscures indications of Army City's features. A low resolution color image from 2004 (probably in late fall), also from Fort Riley, shows greener vegetation in and about the study area over the principal archaeological features of Army City, as defined by the geophysical surveys (see Section 4.6.1). This vegetation marking shows a number of linear and rectilinear patterns that probably correspond in a general way to buried elements of the town (Figure 4.57d). A high resolution color image of the study area obtained from the powered parachute in the spring of 2004 is shown in Figure 4.57e. It reveals very little about the buried town, but indicates the presence of numerous "prairie rings," a form of grass growth on the Plains that also impacts the thermal infrared imaging (discussed below). Finally, the panchromatic band from a QuickBird satellite scene dated April 3, 2001, with 0.6 m spatial resolution, reveals a series of linear and rectilinear anomalies and clear indication of Washington Avenue (Figure 4.57f; the QuickBird imagery is again revisited in Section 4.9). Despite a variety of Army City images, it is apparent that available normal-light aerial or space views do not appear to indicate much about the town's subsurface structure.

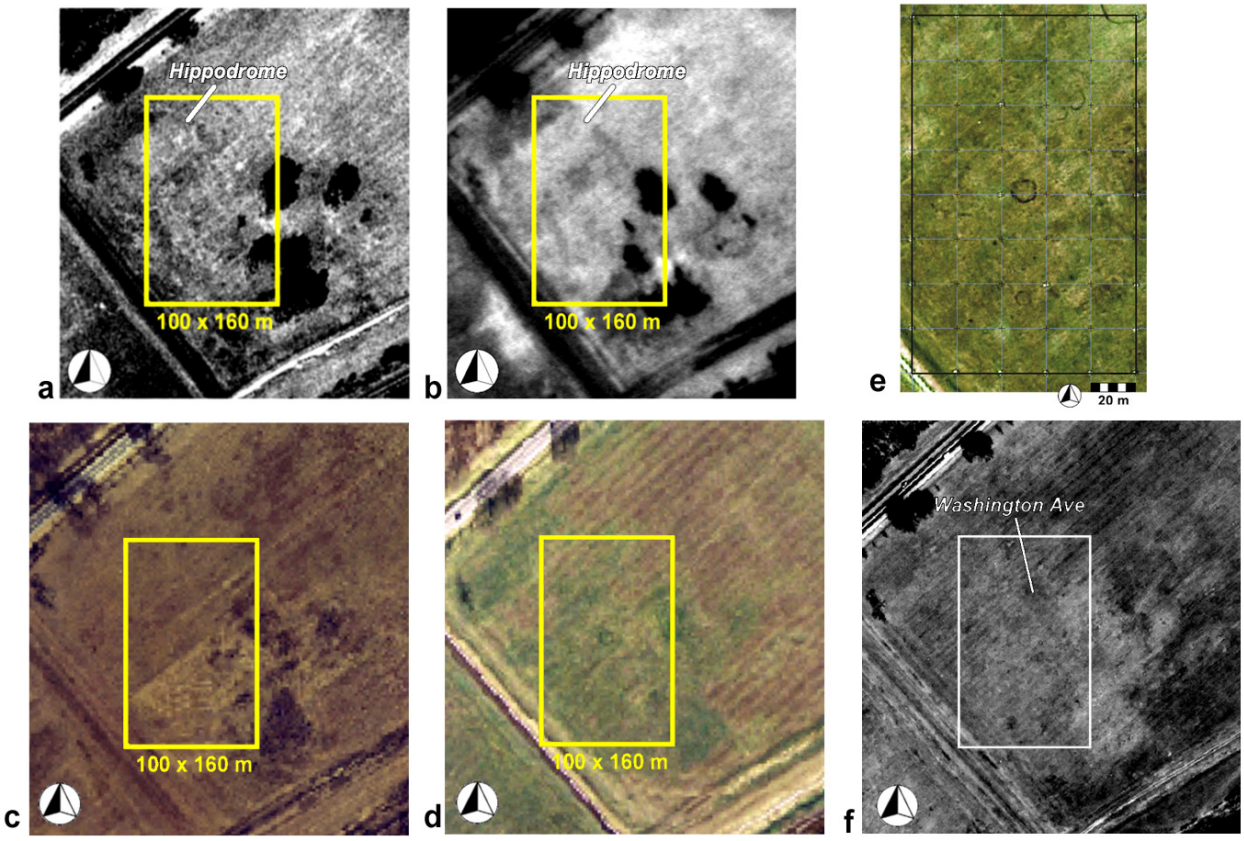


Figure 4.57. *Aerial imagery from Army City. a) NAPP 1991. b) NAPP 1996. c) Low resolution soil survey, 1999. d) Low resolution soil survey 2004. e) High resolution from powered parachute 2004 (note circular "prairie rings"). f) QuickBird panchromatic at 0.6 m resolution, 2001. All images are strongly contrast enhanced and the project study area is defined in each.*

4.8.5. Thermal Infrared Imagery from Army City

Kenneth L. Kvamme, University of Arkansas

The thermal infrared data acquired by powered parachute at Army City were generally informative. One problem was motion blurring endemic to the Palm-IR250 thermal infrared camera. About two-thirds of the videotape was significantly blurred, but the videotapes were painstaking reviewed (at 30 frames per second) allowing extraction of hundreds of unblurred frames from throughout the study area. Several frames acquired at low altitudes yielded exceptional detail. Figure 4.58a shows a number of anomalies that represent building footings significantly warmer than the surrounding soil (white dots). This view was acquired in an evening flight immediately after sunset when dense materials like brick, concrete, or stone yet retain the day's warmth. It illustrates sub-meter detail at a spatial resolution of approximately 6 cm. These data are enhanced through a high-pass filter in Figure 4.58b. Data from the corresponding electrical resistivity survey (Section 4.6.1) indicates similar anomalies, but these data are much coarser with a spatial resolution of only 50 cm (Figure 4.58c). The high resistivity anomalies points to brick, concrete or stone footings, consistent with thermal indications and likely footing materials.

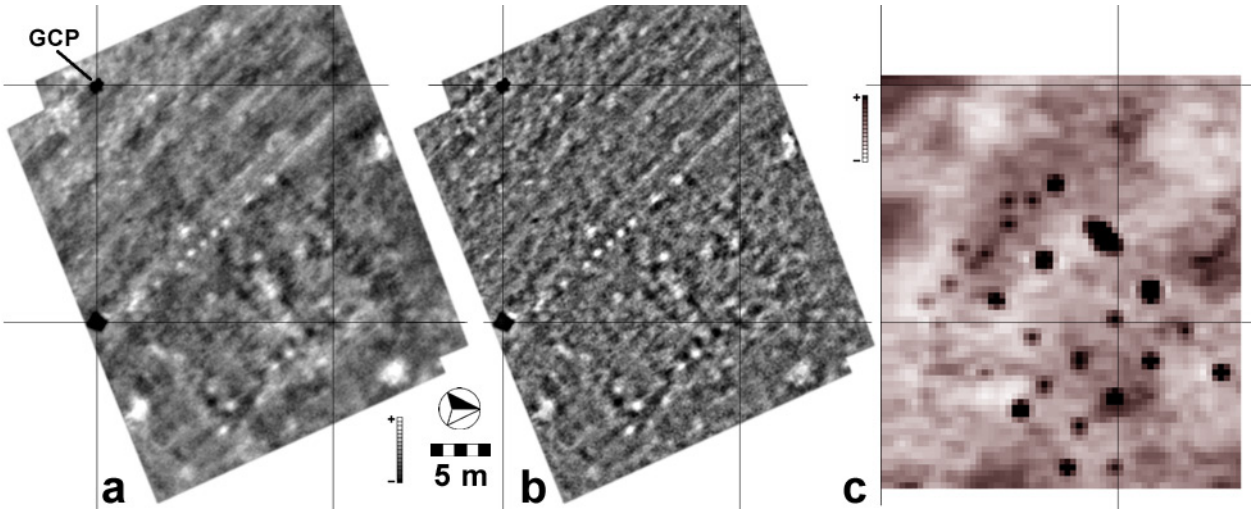


Figure 4.58. *Thermal infrared imagery from Army City showing details of building footings. a) Raw imagery. b) Enhanced through a high-pass filter. c) Corresponding electrical resistivity data.*

Similar detail is shown in Figure 4.59a. This image reveals elements of the Hippodrome, a commercial structure (see Section 3.3.4). A number of prominent walls and building footings are clearly visible (arrows), as are hints of Washington Avenue and General Street. These data, acquired in an early morning flight, show inverted thermal properties with footings and some of the walls appearing cooler than the background. Corresponding resistivity data (Section 4.6.1) are given to clarify findings (Figure 4.59b).

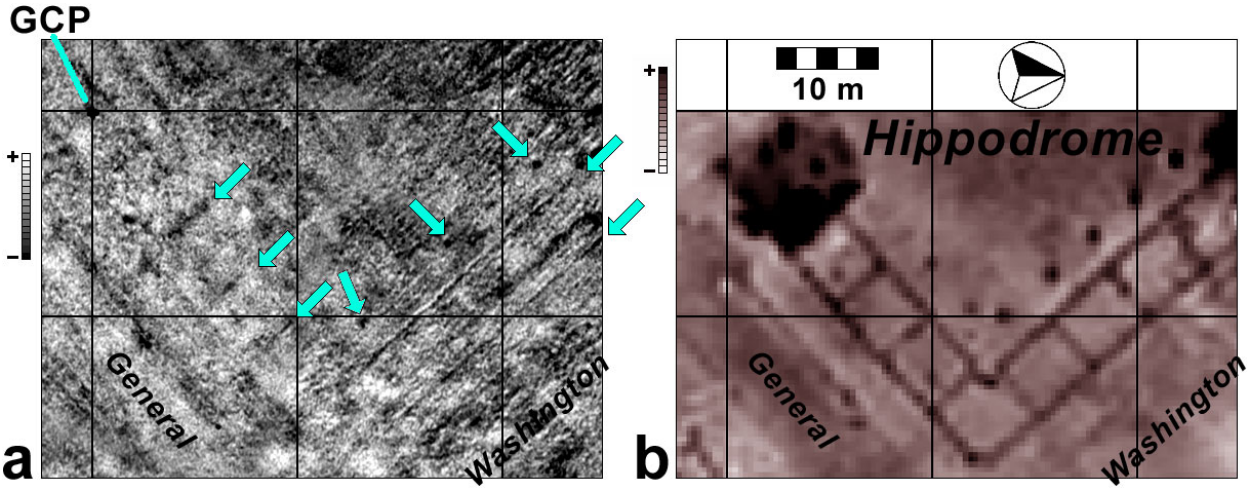


Figure 4.59. *Thermal infrared imagery from Army City showing wall and footing details of the Hippodrome. a) Raw imagery, with arrows pointing to elements of obvious interest. b) Corresponding electrical resistivity data.*

A full study-area-wide composite of mosaic thermal infrared imagery is given in Figure 4.60. The levee ditch, very wet, appears cool while the levee itself is much warmer. Of more significance, numerous anomalies and many important elements of Army City are indicated. Washington Avenue is prominently revealed and General and Colonel Streets can be discerned. Moreover, a number of prominent rectangular and

linear areas illustrate warm or cool anomalies, pointing to likely concrete, brick, or perhaps tiled areas for the former (e.g., floors), and recessed floors, gutters, or perhaps cellars filled with moist sediments for the latter. Also prominent are a number of prairie rings with the outer perimeter of new growth standing tallest and therefore cool (white arrows; see also the color view in Figure 4.57e). This data set is extensively employed in data fusions in Section 5.

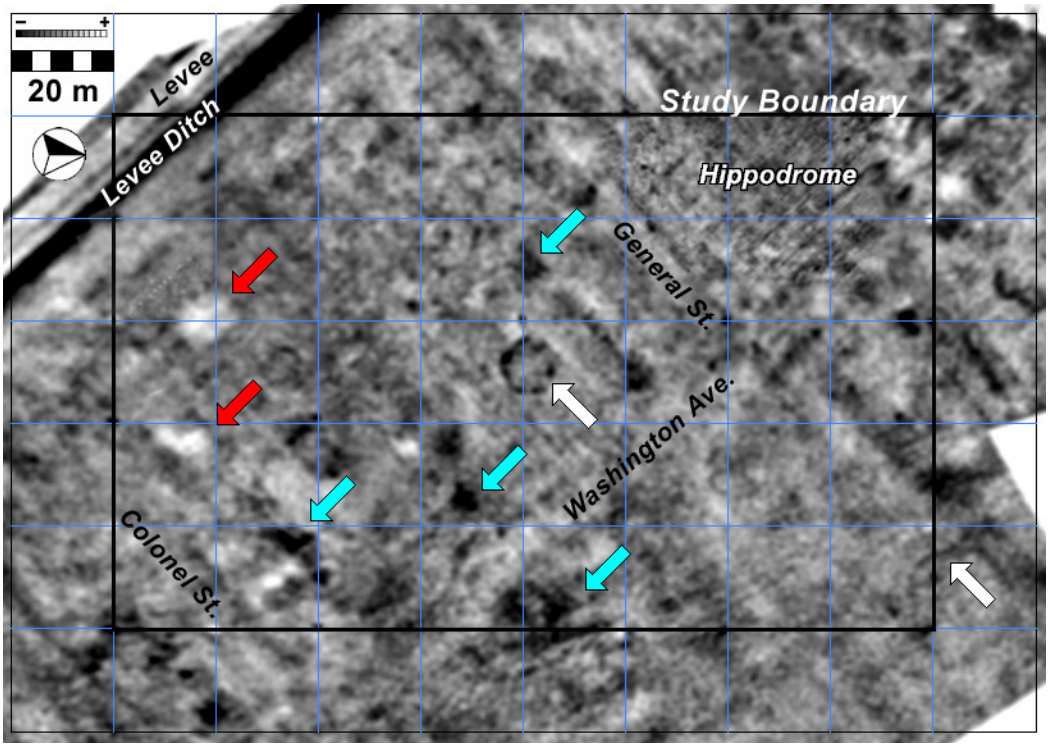


Figure 4.60. *Thermal infrared composite for Army City. Red and blue arrows point to warm and cool anomalies, respectively. White arrows point to "prairie rings."*

4.8.6. Thermal Infrared Results from Pueblo Escondido

Eileen G. Ernenwein, University of Arkansas

Thermal data were collected at Pueblo Escondido in January 2005 using the Raytheon Palm IR 250 thermal imager, which was hand-operated from a small two-seat helicopter, rented for the occasion. The flight was planned so that data could be collected for approximately 25 minutes before and after sunset, but despite prior approval, permission for entering Fort Bliss airspace proved to be difficult once in the air and the flight was delayed for about 20 minutes. The total flight time over the site was approximately 25 minutes most of which was after sunset. During the short flight it was difficult to fly slowly enough to prevent blurring of the imagery and most of the recorded video is unusable due to inherent problems with this instrument. This was the first and only attempt to collect thermal data from this platform, and the instrument proved to be very sensitive to motion—even at the slowest speeds the helicopter pilot was willing to travel.

Several frames from the thermal video were captured as images and used to create a mosaic rectified and registered to the coordinate base of the principal geophysical

survey area (see Section 4.6.2). Twelve metal GCP were placed at grid corners every forty meters, making it possible to rectify the images to the plane coordinate system of the geophysical grid (as in Figure 4.51). The result is very similar to the much of the satellite and aerial imagery investigated by the NASA team below, which mainly show patterns of vegetation. No obvious cultural features are apparent in these data.

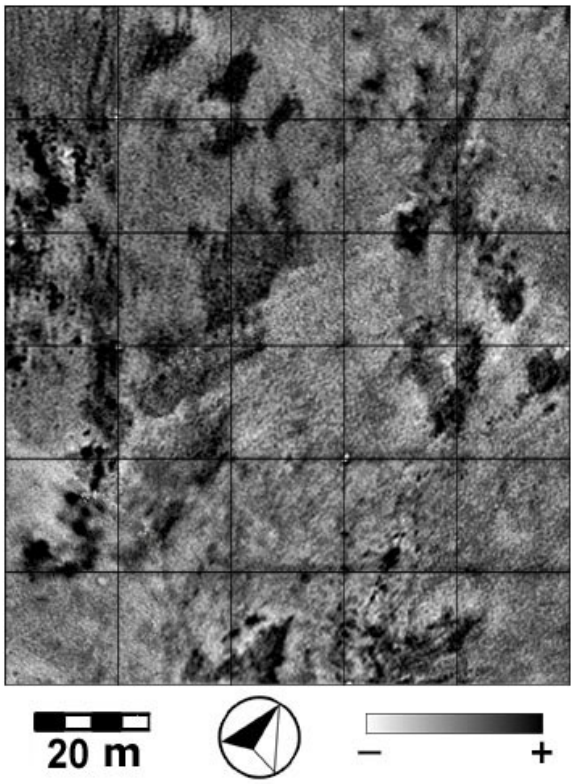


Figure 4.61. *Thermal infrared imagery over principal study block at Pueblo Escondido, acquired with the Raytheon PalmIR250 thermal imager.*

4.8.7. Relationship between Satellite Panchromatic, Aerial Thermal, Conductivity, and GPR data at Pueblo Escondido

Eileen G. Ernenwein, University of Arkansas

Several aerial and satellite data sets were acquired over Pueblo Escondido, including multispectral and panchromatic imagery from the QuickBird satellite, National Aerial Photography Program (NAPP) photos (Figure 4.57), and aerial thermal images acquired with CAST's thermal camera from a helicopter (Figure 4.61). While these data have provided little direct information about Pueblo Escondido, they exhibit interesting relationships with each other and with the ground-based geophysical data. Aerial and satellite methods for archaeological site exploration rely heavily on vegetation, which, when conditions are favorable, can very clearly delineate buried cultural remains such as fortification systems, canals, roads, and prehistoric houses (Wilson 2000). No such patterns are evident in the Pueblo Escondido data, however, probably due to the lack of consistent vegetation cover on the dry desert floor (Figure4.62a).

Although large expanses of bare soil are often a hindrance for prospecting in the visible spectrum, they can be favorable for thermal infrared sensing because the best

opportunity for detecting subsurface features may be over bare soil (Perisset and Tabbagh 1981). This could be true for Pueblo Escondido, but unfortunately the use of a helicopter caused too much motion resulting in blurred imagery and a lack of fine details (Figure 4.61). When compared against a QuickBird satellite panchromatic image of the region, acquired on March 26, 2003 (see also Section 4.9), it is clear that these data sets respond most strongly to the presence and absence of vegetation (Figure 4.62b; the spatial offset between the panchromatic and thermal image is due to small errors in georegistration). It should also be noted that the thermal data were captured some fifteen months after the vegetation had been almost entirely cleared for geophysical survey, and that some differences should likely exist.

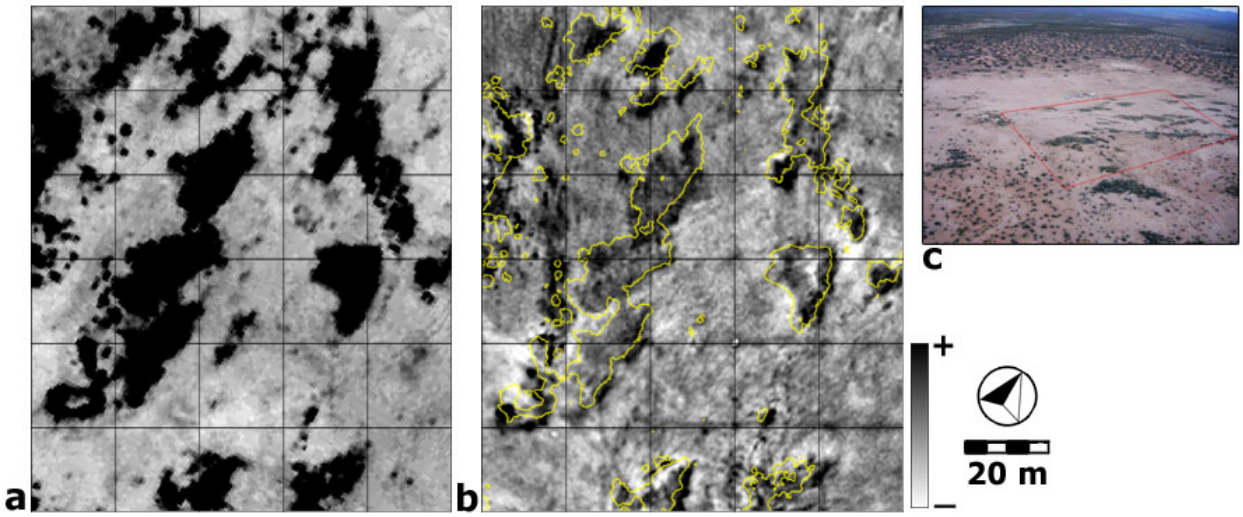


Figure 4.62. *Comparison between satellite and aerial data over the principal Pueblo Escondido study area. a) QuickBird panchromatic image at 0.6 m spatial resolution. b) Aerial thermal image with vegetation based on QuickBird panchromatic data shown with yellow contours. c) Color aerial photograph of the site looking south with approximate location of geophysical survey area shown in red.*

The QuickBird panchromatic and thermal infrared data sets also relate to some of the ground-based geophysical data. Figure 4.63a illustrates the areas of high EM conductivity in black, with vegetated areas indicated by the QuickBird panchromatic band outlined in yellow. Although the relationship is not consistent throughout the entire study area, it is clear that the majority of the vegetated areas are associated with regions of very low conductivity, whereas areas lacking vegetation exhibit high and low conductivity. Vegetation draws moisture from the ground. It may also promote low conductivity by loosening up soil with root systems, thereby increasing porosity and allowing it to dry out more thoroughly. In the right moisture conditions this could make root zones less conductive. The satellite and aerial data have therefore aided the interpretation of the ground-based conductivity data. What the conductivity data are showing remains unclear, but it appears that some anomalies are related to the distribution of vegetation.

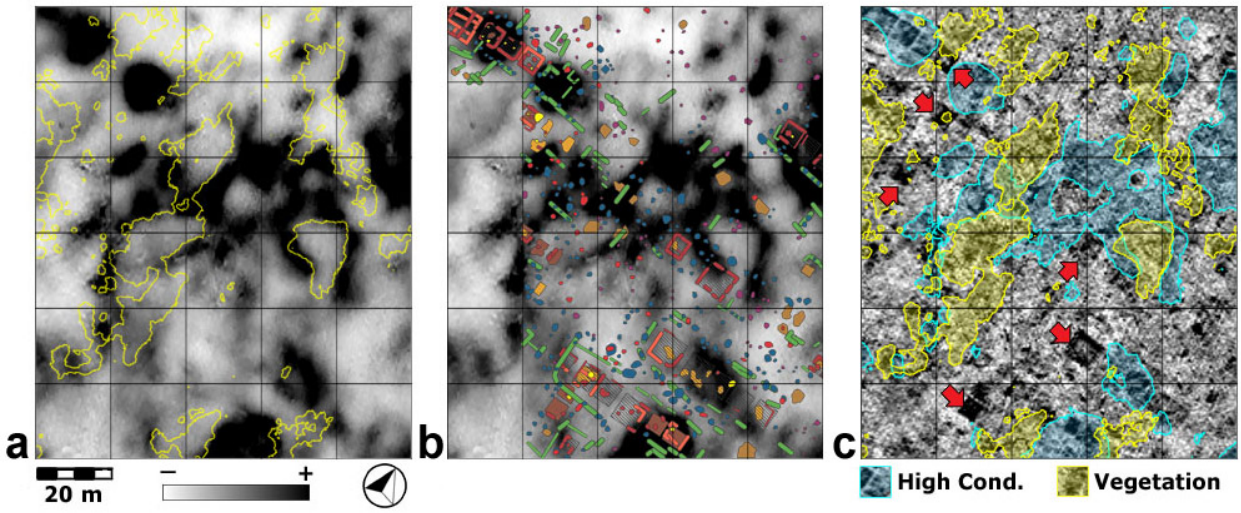


Figure 4.63. *Pueblo Escondido geophysical data and patterns of vegetation. a) Conductivity data with vegetated areas outlined in yellow. b) Distribution of principal anomalies, most revealed by GPR, over conductivity data. c) Zones of high conductivity and vegetation overlaid on time-slice (31-47 cmbs) showing relationships with GPR responses.*

Patterns in the QuickBird, thermal, and conductivity data are also related to the ground-penetrating radar response. When the vectorized interpretations, which are based largely on GPR, are compared to the conductivity image (Figure 4.63b), it is clear that areas of high conductivity contain fewer and less distinct archaeological features. This is more apparent in Figure 4.63c where high conductivity and vegetated areas are shown overlaid on a 31-47 cm deep GPR slice. It is not surprising that areas of high conductivity, which attenuates GPR signals (Conyers 2004), are associated with few GPR reflections. Vegetated areas, generally associated with greater ground disturbance, are also associated with poorer GPR results. In fact, the panchromatic and conductivity data could be used in the future as a means of predicting GPR success.

It is clear that GPR responds much better in areas of low conductivity lacking vegetation, but the exact causes are still unknown. Vegetated regions might be associated with poor preservation, as could high conductivity areas. Alternatively, preservation may be similar across the site, but imaging with GPR is simply compromised by the presence of vegetation and high conductivity. These questions were not addressed during the archaeological excavation program, but no major differences in preservation were found to be associated with vegetation or conductivity. Although aerial and satellite data do not appear to be highly successful methods of locating archaeological features at Pueblo Escondido, they have proven useful when used together with ground-based geophysics. The QuickBird panchromatic image shows that the thermal data responded primarily to vegetation cover, which is also associated with low conductivity. In addition, the GPR data are much improved in areas of low conductivity and little or no vegetation, suggesting that aerial images and conductivity could help predict the success and reliability of GPR over a survey area.

4.9. AERIAL AND SPACE MULTI-SENSOR REMOTE SENSING BY NASA

Thomas L. Sever and Burgess F. Howell, NASA Marshall Space Flight Center

4.9.1. Introduction

Kenneth L. Kvamme, University of Arkansas

This section summarizes aerial and space-based remote sensing activities conducted by the NASA Marshall Space Flight Center at the SERDP project sites. The NASA team performed an extensive array of aerial remote sensing and additionally conducted research using high resolution QuickBird satellite imagery. Sever and Howell (2005) present this work in a summary report. This team chose to perform their analyses "with little a priori knowledge of specific cultural features on the ground in order to determine the ability of aerial remote sensing to locate and map unknown archeological features without the benefit of ground verification." Because it offers a "blind" appraisal of the SERDP project sites, the following material is very different from the content of foregoing sections. Much focus and many interpretations were consequently made *outside* of the study boundaries of the geophysical surveys, for example. Some of this interesting material cannot be presented here, where focus continues on the specific project areas and multidimensional remote sensing findings within them. Project study areas have been superimposed on the NASA imagery to maintain that focus in this report. Comparisons of these findings should be made against geophysical and aerial remote sensing materials given in Sections 4.6 and 4.8 to aid in the interpretation of these results. In particular, some of the remote sensing interpretations given at the various sites point to disturbed brushy ground where trees were once located at Army City, to known pipelines at Kasita Town, and geophysical survey grid lines (that impacted plant growth from repeated walking) at Silver Bluff and Kasita Town.

The following sections describe data sources, modes of acquisition, processing and enhancement methods, and analytical techniques. Several problems encountered while collecting and processing the data are enumerated, as well as lessons learned that may be applied to future projects of this nature.

4.9.2. Sensor Instrumentation

4.9.2.1. QuickBird

QuickBird is a high-resolution commercial earth observation satellite owned by DigitalGlobe, launched in 2001. It collects panchromatic (black and white) imagery at 60-70 centimeter resolution and multispectral imagery at 2.4- and 2.8-meter resolutions. Bands and band widths are given in Table 4.5. The satellite orbits the earth every 93.4 minutes at an altitude of 450 km in a 98 degree sun synchronous inclination. It acquires data over an area of interest either in a single area (16.5 km by 16.5 km) or in a strip (16.5 km by 165 km) and revisits the same location on the earth every 1 to 3.5 days, depending on latitude at 70 cm resolution.

4.9.2.2. Raytheon PalmIR 250

The PalmIR 250 is an inexpensive, lightweight portable thermal unit that was mounted on a fixed wing aircraft flying at high altitude for complete coverage of the study area and environs in single scenes. It is a single, broad band (7-14 microns) thermal system that detects slight temperature differences between objects and people in its field

of view and uses this information to create real-time "thermal landscapes" of the area on display. It uses a 320 X 240 barium strontium titanate detector with a spatial resolution of 1.0 mrad and a thermal sensitivity of less than 0.10 at 30 degrees Centigrade. Thermal data are direct results of target and temperature. Due to the low altitude at which the PalmIR 250 data were acquired, the atmospheric component was negligible and no atmospheric corrections were made. Emissivity, temperature, and reflectance are therefore features that influence the radiation detected.

Table 4.5. *QuickBird bands and band widths.*

Channel 1	pan	450-900 nm
Channel 2	blue	455-516 nm
Channel 3	green	506-595 nm
Channel 4	red	630-690 nm
Channel 5	NIR	760-900 nm

4.9.2.3. DuncanTech MS4100 Digital Multispectral Camera

The DuncanTech MS4100 camera is a digital, progressive scan, multispectral camera mounted on a fixed wing aircraft. It employs a color separating prism and three imaging channels that allow simultaneous image acquisition in 3-5 spectral bands through a common aperture. Image sensors are charge coupled device (CCD) array sensors (Table 4.6).

Table 4.6. *DuncanTech MS4100 bands and band widths.*

Channel 1	green	520-590 nm
Channel 2	red	630-690 nm
Channel 3	NIR	760-900 nm

4.9.2.4. Advanced Thermal and Land Applications Sensor (ATLAS)

Six thermal bands from ATLAS were analyzed over Pueblo Escondido at Fort Bliss, Texas. The data were provided by personnel from Ft. Bliss who had previously acquired extensive ATLAS data for interdisciplinary research at 5 m spatial resolution. Only a subset of the data, the area over Escondido Pueblo, was provided to this project. The ATLAS airborne sensor acquires high spatial resolution multispectral and thermal infrared data and is flown on board a Lear 23 jet aircraft operated by the NASA Stennis Space Center. The ATLAS is a 15-channel multispectral scanner that basically incorporates the bandwidths of the Landsat TM (along with several additional channels) and six thermal IR channels similar to that available on the airborne Thermal Infrared Multispectral Scanner (TIMS) sensor (Table 4.7). Only the thermal channels, bands 10-15, were provided and used in this study.

Table 4.7. *ATLAS sensor system specifications.*

Channel NUMBER	Bandwidth Limits (µm)	NER mW/cm^{-2} µm	NE T ^{o}C	MTF @ 2 mrad	Cooling
1	0.45-0.52	<0.008	N/A	0.5	Ambient
2	0.52-0.60	<0.004	N/A	0.5	Ambient
3	0.60-0.63	<0.006	N/A	0.5	Ambient
4	0.63-0.69	<0.004	N/A	0.5	Ambient
5	0.69-0.76	<0.004	N/A	0.5	Ambient
6	0.76-0.90	<0.005	N/A	0.5	Ambient
7	1.55-1.75	<0.05	N/A	0.5	77 ^{o}K
8	2.08-2.35	<0.05	N/A	0.5	77 ^{o}K
9	3.35-4.20	N/A	<0.3	0.5	77 ^{o}K
10	8.20-8.60	N/A	<0.2	0.5	77 ^{o}K
11	8.60-9.00	N/A	<0.2	0.5	77 ^{o}K
12	9.00-9.40	N/A	<0.2	0.5	77 ^{o}K
13	9.60-10.2	N/A	<0.2	0.5	77 ^{o}K
14	10.2-11.2	N/A	<0.2	0.5	77 ^{o}K
15	11.2-12.2	N/A	<0.3	0.5	77 ^{o}K

4.9.3. Image Analysis Software

Image processing software utilized for this study included ERDAS Imagine (1999), RSI-ENVI, and Z/I Imaging Image Analyst. ERDAS Imagine is the industry standard for processing remotely sensed data and includes a powerful graphical modeling capability to develop customized models and streamline routine tasks. RSI-ENVI is a competitive image processing product that was utilized for specific tasks that were problematic in Imagine. For example, RSI offers improved capabilities for post classification and spectral mapping. Also, ENVI color tables are generally superior to ERDAS. Image Analyst was utilized for digitizing areas of interest. It is particularly well suited to this task because of its flexible and powerful image display tools that allow interactive real-time modification of image display characteristics. ELAS is a sophisticated image processing and remote sensing package that allows the user great flexibility in modifying the operations and parameters at a very fundamental level.

4.9.4. Processing Techniques

4.9.4.1. Georeferencing

All data sets used in site analysis were projected (or reprojected, where necessary) to a common workspace on a per site basis using the Universal Transverse Mercator system. Since QuickBird data were common across all sites, each of those scenes was arbitrarily chosen as being in the “correct” orientation, and every other data set of interest was registered to its respective QuickBird analog.

4.9.4.2. Resolution enhancement

In order to enhance the spatial resolution of the multispectral QuickBird data used in this project, a data fusion process known as pan sharpening was performed. In pan sharpening one of several techniques is employed to generate new data sets which combine the high frequency spatial domain data of the panchromatic band with the higher information content, but lower spatial resolution data of the multispectral bands. For the data in this study, we used a process known as a Principal Components Transformation (PCT). PCT transforms the multispectral data from spectral space into a feature space where one of the resulting bands is correlated with the high resolution data. The high resolution band is then substituted for the correlated band and a reverse transform applied to translate the bands back to spectral space. In the PCT, information content common across the multispectral bands is mapped to a single output data set called the Principal Components Image (PCI). The content of this image is typically highly correlated to brightness or intensity (Jensen 1996; Chavez et al. 1991). The high resolution QuickBird panchromatic data are contrast matched to the PCI, then substituted for the PCI in an inverse PCT operation that transforms the data back into spectral space (Schott 1997).

4.9.4.3. Image stacking

A simple but effective technique to examine the data content of dissimilar data sets is to use them as constituents of an image stack. In much the same way that a GIS system allows examining multiple database layers at a common location, an image stack allows computational and visual analysis of multiple spectral datasets over a common area. We created image stacks for each study area utilizing input bands from 2 or more sensors. In each case, the site's QuickBird s scene was chosen as the "master" data set, with the other data sets being resampled such that their pixels were dimensionally identical to, and spatially coincident with, their QuickBird counterparts.

4.9.4.4. Band ratioing

Over most sites, one or more synthetic data sets were produced by ratioing pairs of image bands. In most instances, the purpose was to generate a normalized difference vegetation index for inclusion in an image stack. Where coincident Duncan and QuickBird data were available, a ratio of the NIR bands from those data sets highlighted changes in vegetation vigor and density between the two data acquisition dates.

4.9.4.5. Filtering

For the thermal data, the appearance of a target varies from image to image across a sequence of frames. When a sequence is put together as a mosaic the variation is manifested as noise that obscures the target signal. To reduce the impact of this noise, a simple 3x3 low pass convolution kernel was passed over the mosaic of thermal data. To regain lost detail from this operation, the output of the low pass kernel was subjected to a Laplacian edge enhancement filter. Those filtered data were then used as final input for stacking, ratioing, and other processing.

4.9.4.6. Principal components analysis (ATLAS thermal)

Each of the ATLAS thermal channels was reviewed individually and various band combinations created in an attempt to detect the ruins of Pueblo Escondido. In addition,

principal component analysis was performed to circumvent noise problems in the data. In this technique a set of axes is selected so that the maximum amount of a data set's variation is accounted for by a minimum number of perpendicular axes. The first principal component can be equated to target brightness and constitutes the majority of the data, while the remaining components contain logarithmically decreasing proportions of the data. Principal components analysis provided a crisp image of the study area, as noted in the scrub vegetation and other surface features, but did not reveal walls, rooms, or other features associated with the pueblo structure. One explanation is that the data was acquired at mid-morning when thermal temperatures are balancing rather than at thermal maximum (solar noon) or minimum (pre-dawn.). Although there might be concern that the 5 m resolution of the data precluded the detection of these feature, this is not the case. Five meter aerial thermal data was successfully used at pueblos in Chaco Canyon, New Mexico, which did in fact reveal walls, trash middens, gates, and emerging and exiting roads (Sever 1990; 1983).

4.9.5. Data Analysis

Processed data sets for each site were interpreted by a variety of manual and automated means. Multispectral datasets were subjected to supervised and unsupervised classifications based on spectral characteristics of component bands. Individual bands were "stretched" (contrast enhanced) by density slicing (application of a continuous color table to a grayscale image based on the relative magnitude of constituent pixel values), thresholding (application of a binary color table to a grayscale image based on the relative magnitude of constituent pixels), and stretching (remapping some or all of the range of pixel values to increase the apparent brightness difference between consecutive values). Contrast stretching techniques included simple linear (remapping some or all of an image's constituent pixel values—chosen by specific values—across the dark-to-bright range of a display system), linear percent (remapping a portion of an image's constituent pixel values—chosen as a percentage of all possible values—across the dark-to-bright range of a display system), and linear piecewise (remapping various segments of constituent pixel values to several non-linear and/or non-contiguous segments of the range of a display system).

Because stretch operations are performed utilizing parametric statistics based on a sample of the total pixels for a band or image, the output appearance can be varied almost infinitely by changing the input sample. In particular, sampling a specific spatial extent on the ground causes the stretched output to enhance small differences across that area. Differences outside that area are minimized to the extent that, when a very small or spectrally homogeneous area is sampled, gross features may appear highly generalized and monotone. Both highly enhanced and highly generalized portions of a stretched image can be used to extract information about a scene. Features of interest within the various data sets were noted by creating a set of overlying bounding or delineating vectors using Bentley MicroStation CAD software. Georeferencing information in the form of UTM and geographic coordinates was attached to those vectors by Z/I Imaging Coordinate System Operations. The vectors were then exported as ESRI shapefiles for incorporation with the project overall GIS.

4.9.6. Features of Interest

In general, other than the geographic coordinates of a central point for each site, all data were acquired, processed, and analyzed without benefit of *a priori* knowledge, with two notable exceptions. At Army City, placement of positional reference markers was accomplished with the assistance of other team members, and utilizing a grid originally devised during the course of previous investigation. At Silver Bluff, similar markers were placed at the site alongside and obvious archeological excavations already in progress. Primarily because of this lack of an existing knowledge base upon which to base specific conjecture, most instances of significant or interesting features found in the imagery of the four study sites are noted simply as "features" or "anomalies." Only in cases where feature characteristics make identification obvious do we label targets as to particular type.

4.9.6.1. Army City

A full QuickBird scene of the region around Army City, including much of Fort Riley, is illustrated in Figure 4.64. It is presented in a true color format, with bands 3,2,1 assigned to red, green, and blue, respectively, The general region of interest is illustrated in the yellow box.

Figure 4.64. *Full scene, true color QuickBird image, band 3,2,1 (RGB) showing the general area of study at Army City in yellow.*

Composite imagery is illustrated in Figure 4.65a-d. It shows QuickBird band 4 (NIR) on the red channel, PalmIR thermal on the green channel, and QuickBird band 2

(green) on the blue channel. Figure 4.65b highlights some of the major bright anomalies (in yellow). In addition, much of the area in the left half of the image (particularly in the geophysical study area, is shown to be disturbed ground (in purple). That purple signature extends throughout the image. This scene emphasizes many anomalies outside the geophysical survey area. A close up view of five of the highlighted anomalies is given in Figure 4.65c,d, enhanced to show greater detail.

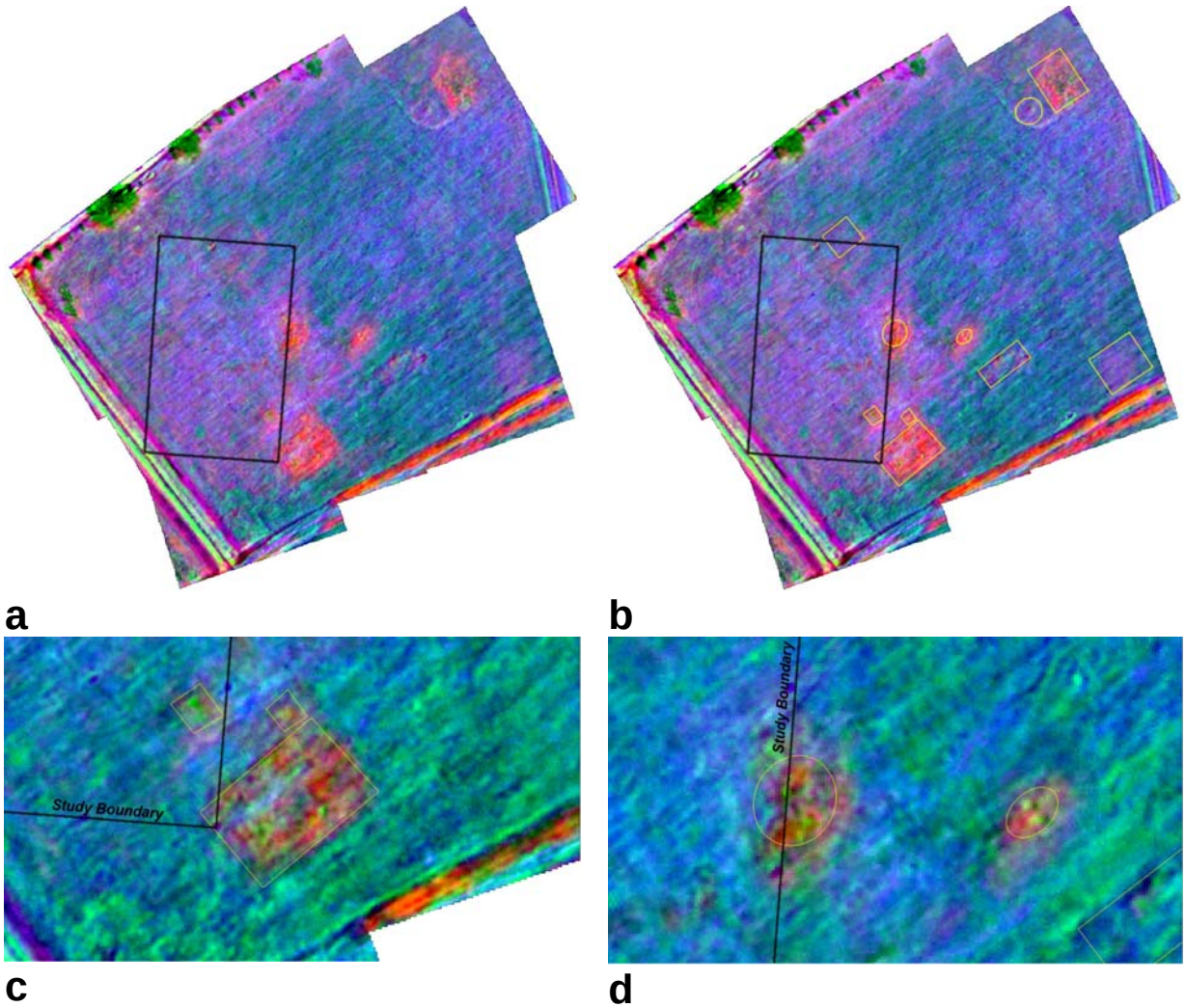


Figure 4.65. *Composite Army City imagery showing QuickBird band 4 (NIR) as red, PalmIR thermal as green, and QuickBird band 2 (green) as blue. a) View of processed region with b) bright anomalies indicated in yellow. c,d) Enhancements focused on specific anomalies. The focal study area of 100 x 160 m is outlined in black.*

The PalmIR thermal infrared data show temperature variations throughout the region (Figure 4.66a). (These data were acquired at relatively high altitude from a fixed wing aircraft, compared to low-altitude results given in Figure 4.60.) The anomalies from Figure 4.65 have been overlaid onto the image. Several small, dark dots in the image are from the thermal reflectance of the aluminum markers used for georeferencing. The data of Figure 4.65 are again illustrated in Figure 4.66b, but with a different enhancement.

Many linear and surface anomalies are superimposed on this image. The linear features represented by red are speculated to be of recent origin.

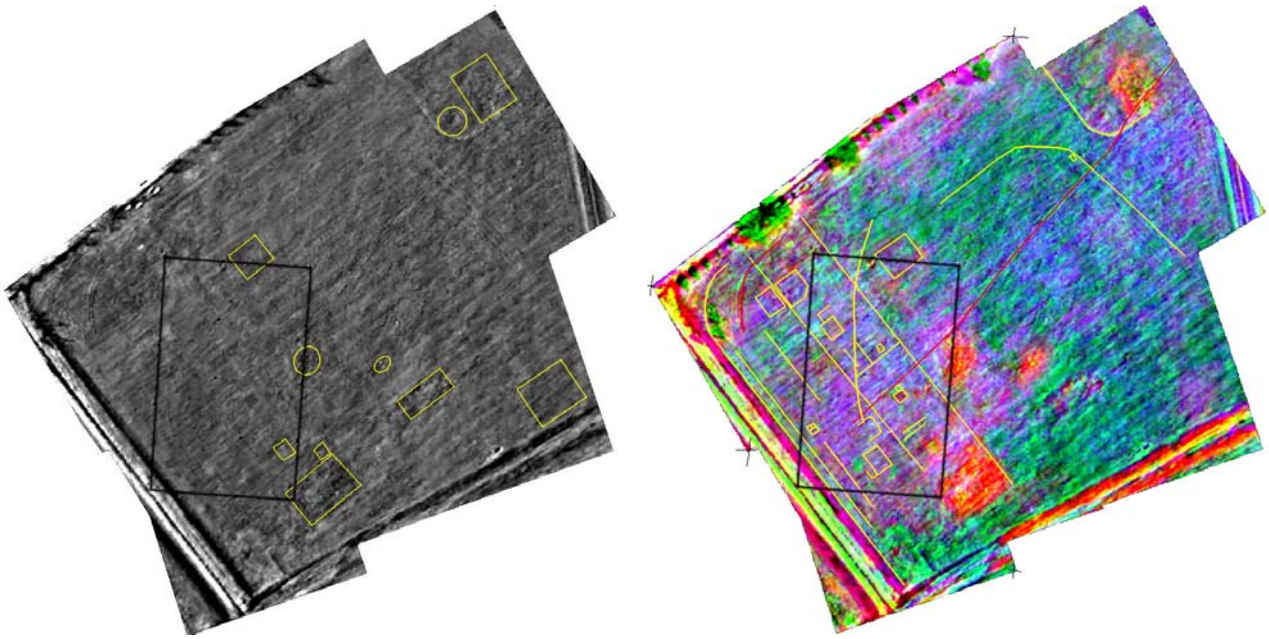


Figure 4.66. *Army City imagery. a) PalmIR thermal temperature variations with anomalies superimposed. b) Enhancement of previous RGB imagery showing linear and surface anomalies. The focal study area of 100 x 160 m is outlined in black.*

4.9.6.2. Pueblo Escondido

The full, true color, QuickBird scene is illustrated in Figure 4.67. This image shows band 3,2,1 as red, green, and blue, respectively. The general area of study is given in yellow.

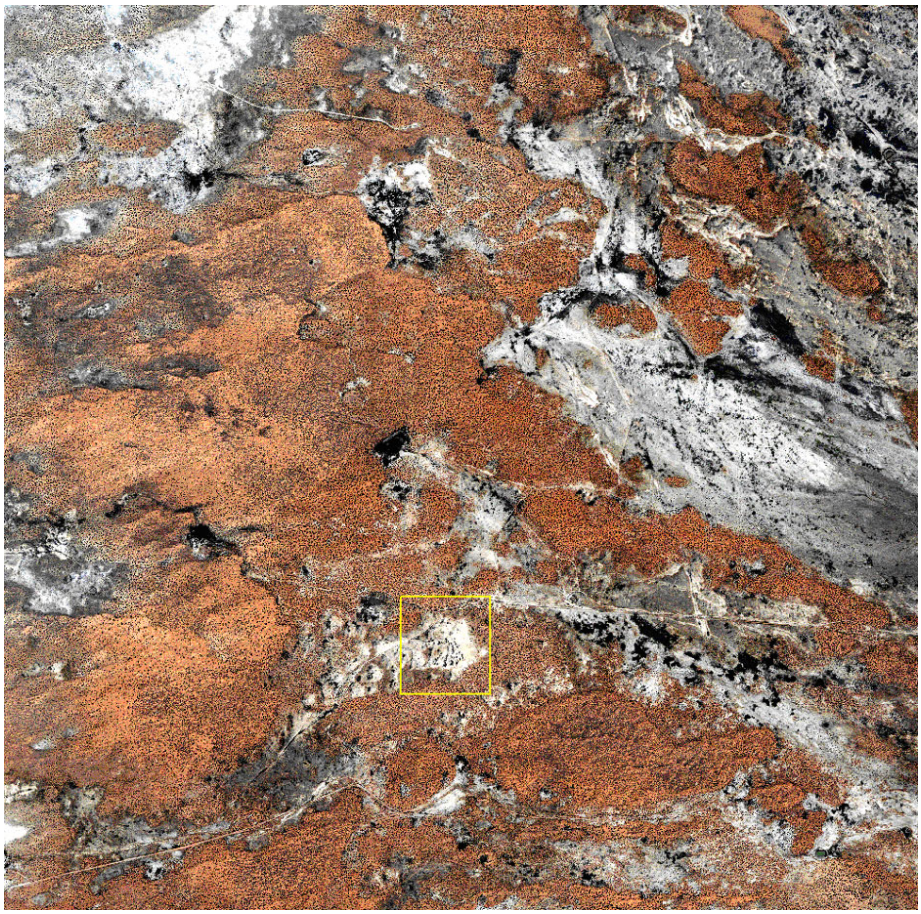


Figure 4.67. *Full scene, true color QuickBird image of the Pueblo Escondido region, using bands 3,2,1 (as RGB). The general area of study is given in yellow.*

A pan sharpened and contrast enhanced QuickBird image is given in Figure 4.68a. This image shows bands 4,2,1 as red, green, and blue (RGB). Unfortunately, no QuickBird data band combinations, processing methods, or enhancement techniques were able to indicate any probable archeological features at this site. This circumstance is apparently true of the other forms of remote sensing data. A contrast enhanced first principal component image of ATLAS thermal data (bands 10-15) is given in Figure 4.68b; a contrast enhanced RGB image of ATLAS thermal principal components 1, 2, and 3 is illustrated in Figure 4.68c; a contrast enhanced RGB image of ATLAS thermal bands 15, 13, and 11 is portrayed in Figure 4.68d. No evidence of cultural features was seen in these data sets.

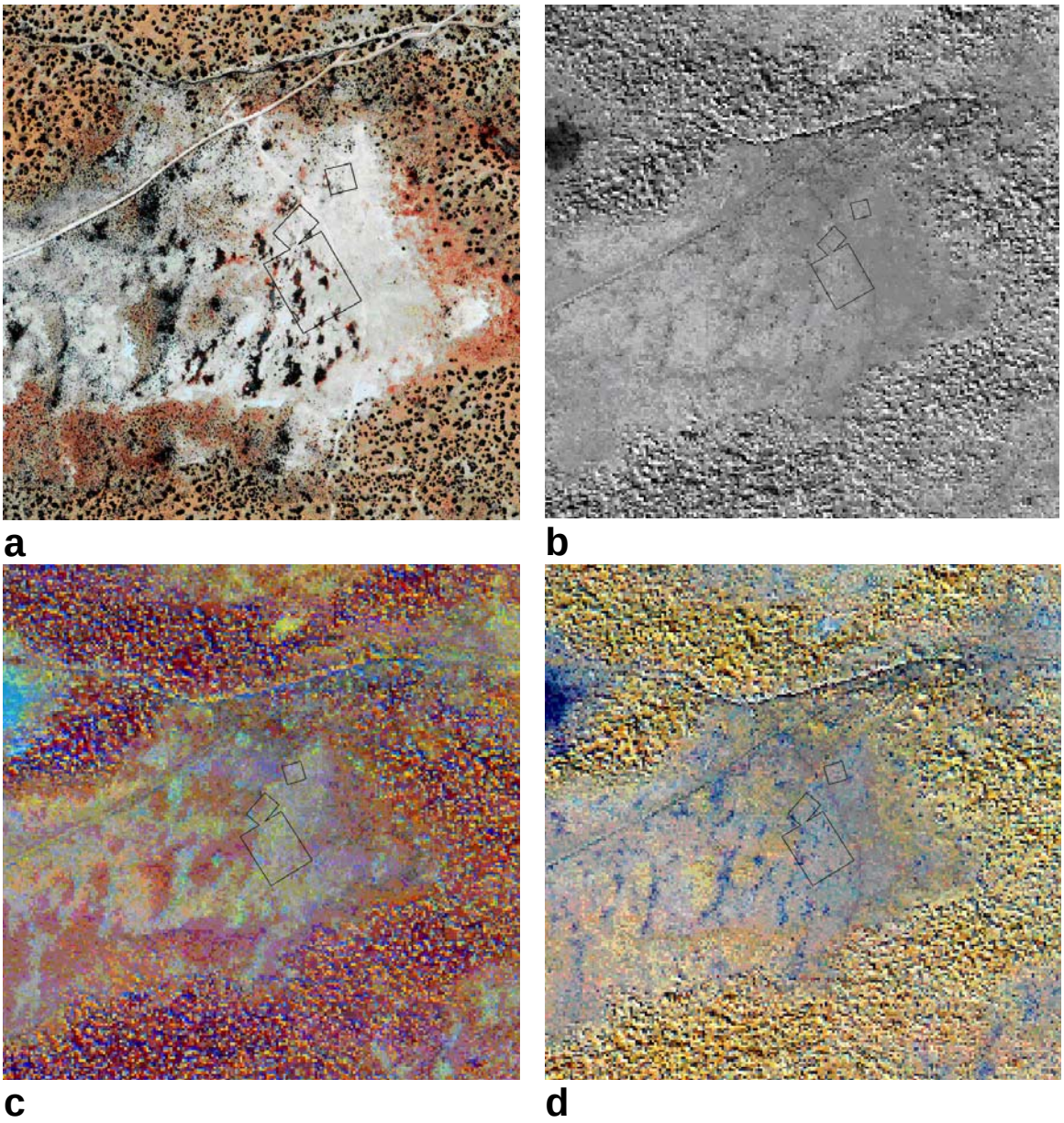


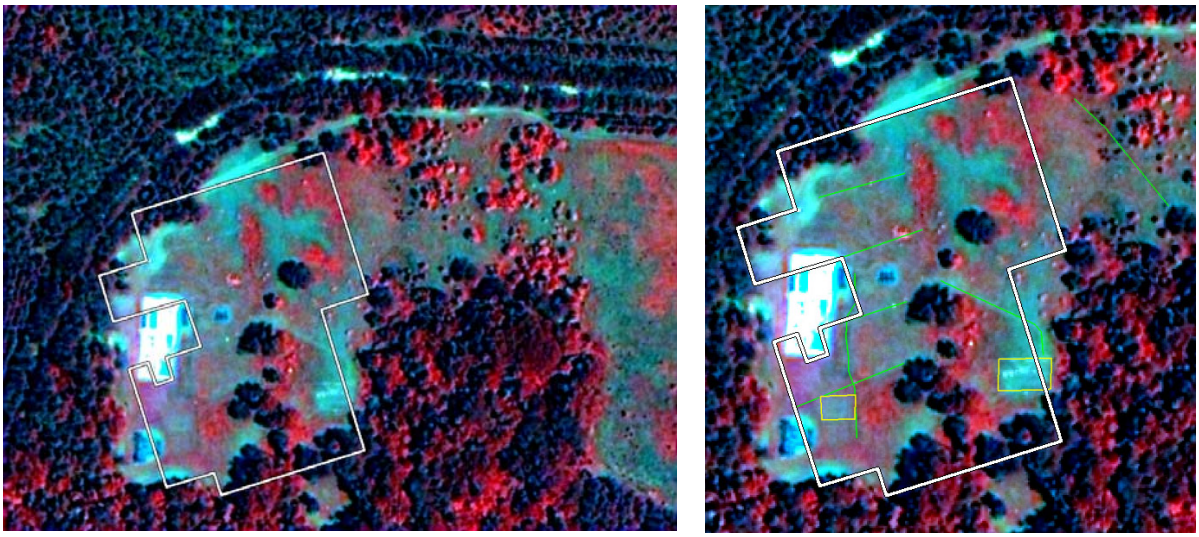
Figure 4.68. *Enhanced imagery of the Pueblo Escondido study region. a) Pan sharpened and contrast enhanced QuickBird image, bands 4,2,1 (RGB). b) Contrast enhanced first principal component image of ATLAS thermal data (bands 10-15). c) Contrast enhanced RGB image of ATLAS thermal principal components 1, 2, and 3. d) Contrast enhanced RGB image of ATLAS thermal bands 15, 13, and 11. The project study areas are outlined in black with the largest rectangular space measuring 100 x 120 m.*

4.9.6.3. Silver Bluff

The full QuickBird scene, a true color image showing bands 3, 2, and 1 as red, green, and blue, respectively, is given in Figure 4.69. The general area of study in outlined in yellow.

Figure 4.69. *Full scene, true color QuickBird image, bands 3, 2, 1 (RGB), showing the general area of study in yellow.*

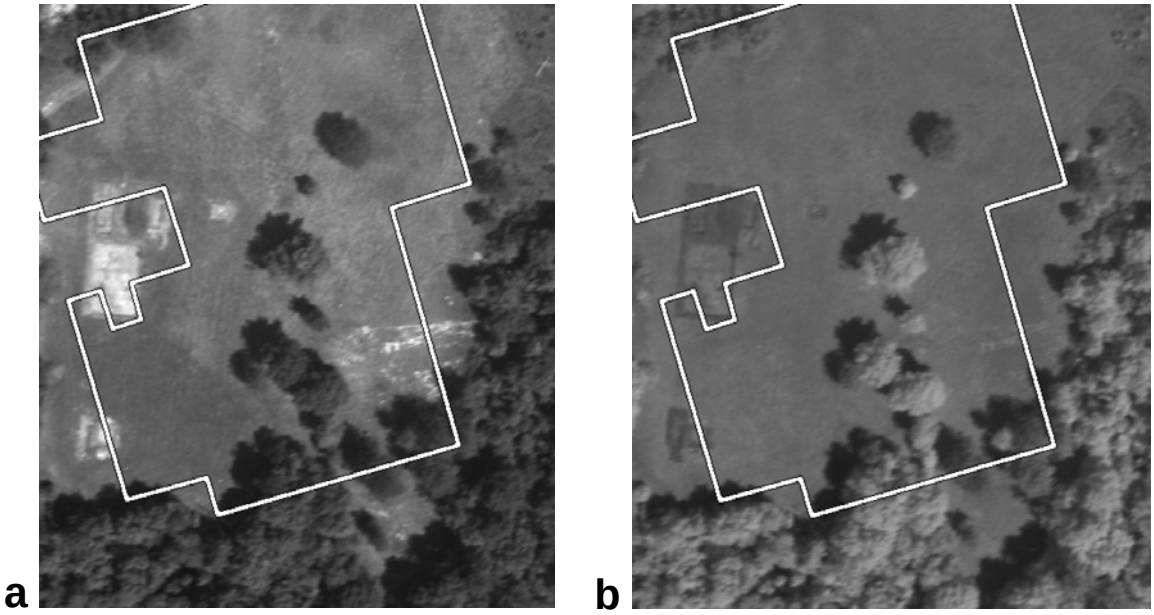
The QuickBird multispectral bands were pan sharpened and contrast enhanced. Bands 4, 2, and 1 are illustrated in Figure 4.70a as an RGB color composite with current excavations visible at lower left (in white and cyan). Interpretations are given in Figure 4.70b. Linear features are outlined in green and two anomalous rectangular areas are defined in yellow. Other anomalies appear in the field as bright red, which are related to vegetation differences.



a b

Figure 4.70. *a) QuickBird image, bands 4,2,1 (RGB) with current excavations at lower left (in white and cyan). b) Linear anomalies are defined in green and two rectangular anomalies in yellow. The project study area is outlined in white.*

The DuncanTech MS4100 camera yielded several clear images of the study area. Band 2 (green) shows excavated areas toward the left of the image and anomalous areas in the clearing near the center of the image (Figure 4.71a). Band 3 (near infrared) shows the enhanced capability of the NIR to detect linear features within the excavations, which may be related to excavation depth or the location of previous structures.



a b

Figure 4.71. *Duncan multispectral data. a) Duncan image band 2 (green). b) Duncan image band 3 (NIR). The project study area is outlined in white.*

4.9.6.4. Kasita Town

The full QuickBird scene, a true color image showing bands 3, 2, and 1 as red, green, and blue, respectively, is given in Figure 4.72. The general area of study in outlined in yellow.

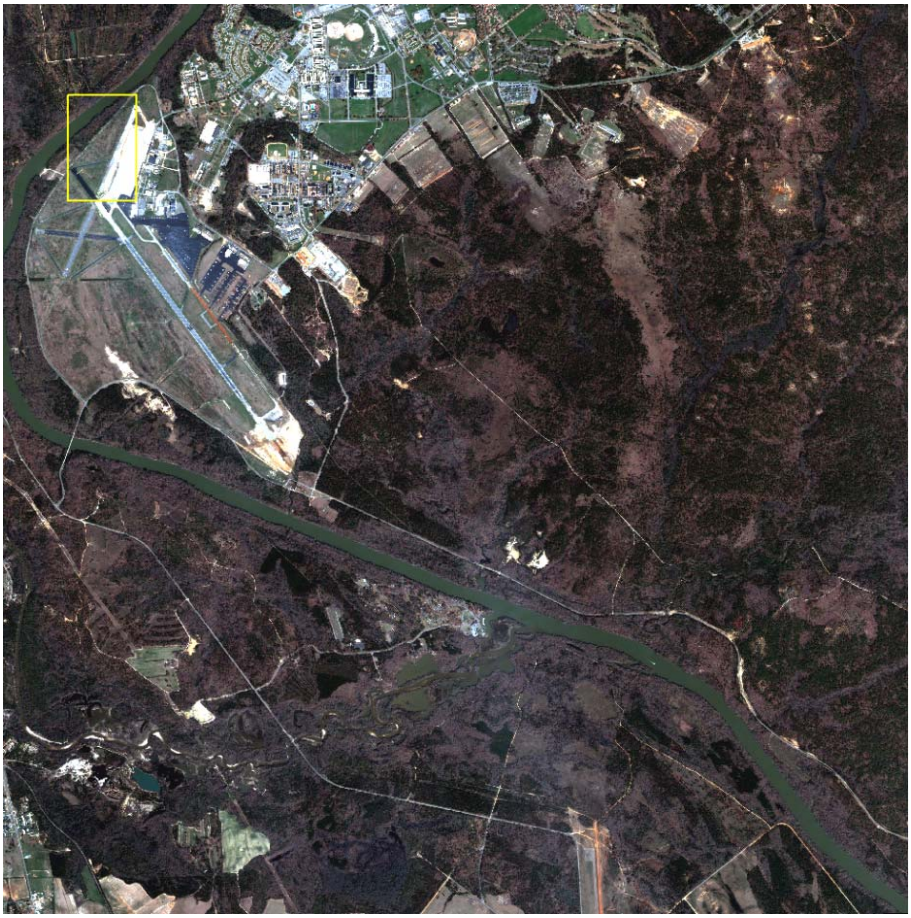


Figure 4.72. *Full scene, true color QuickBird image, band 3,2,1 (RGB) showing the general area of study in yellow.*

At Kasita Town the data were pan sharpened and contrast enhanced. QuickBird bands 4, 2, and 1 were assigned red, green, and blue (Figure 4.73). The field is bounded by a runway to the east (in cyan), and a roadway (cyan with truck convoy) to the west. Two major influences complicated image interpretation: artifacts from mowing and disturbances attributable to current and past runway construction and maintenance.

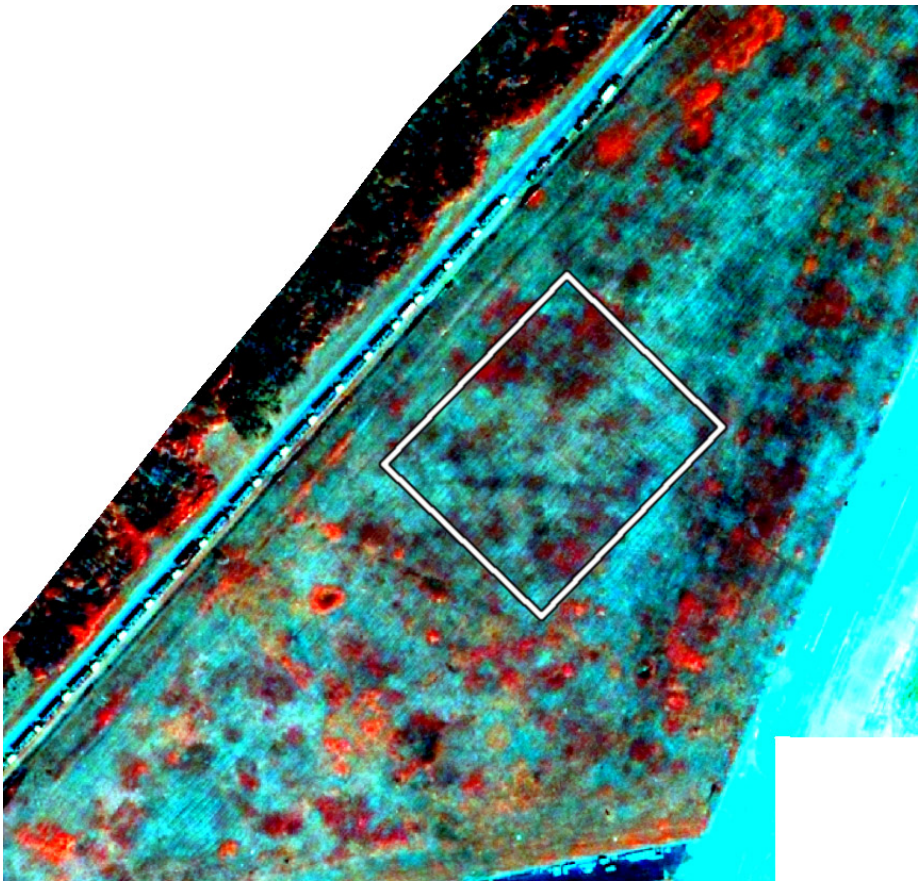


Figure 4.73. *Pan sharpened and contrast enhanced QuickBird image with bands 4, 2, and 1 as RGB. The project study area is outlined in white.*

Several interpretations are offered for these data (Figure 4.74). Multiple linear features and surface anomalies are highlighted in yellow. A variety of other unmarked anomalies are obvious in the image. The longest linear feature extending generally east-west through the scene is a drainage feature. Evidence of this feature continues westward to the river (white area to far left). Many of the anomalies are probably attributable to the ongoing runway construction and maintenance.

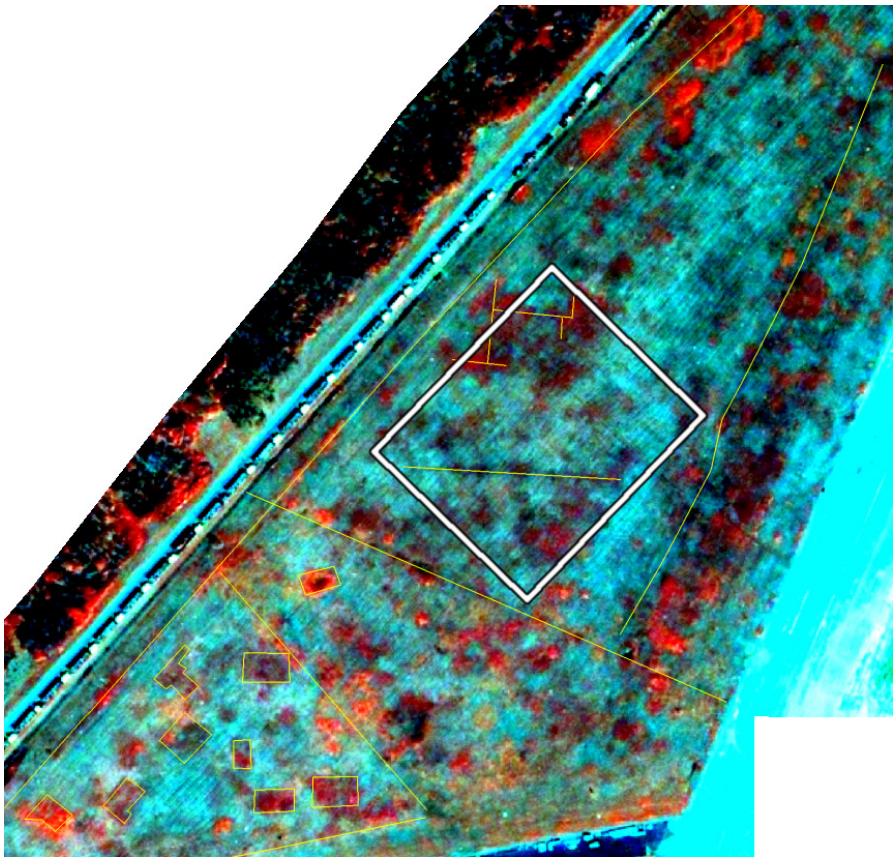
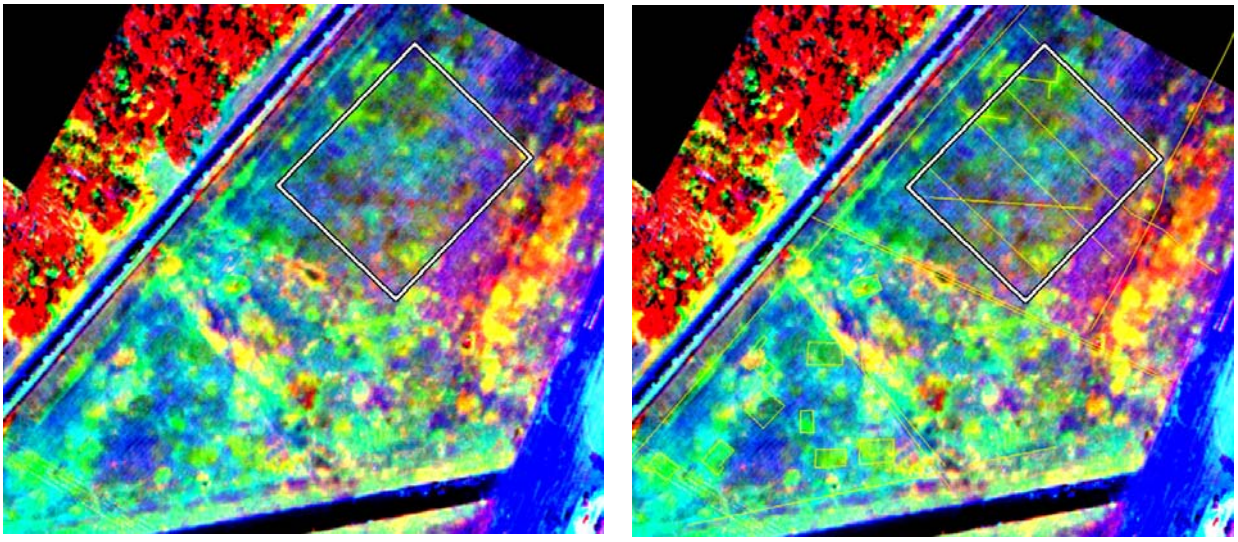


Figure 4.74. *Pan sharpened and contrast enhanced QuickBird image with bands 4, 2, and 1 as RGB. Interpretations are given in yellow. The project study area is outlined in white.*

The DuncanTech MS4100 camera yielded several images of the study area. In Figure 4.75a. a composite image composed of Duncan band 3 (IR) as the red channel, QuickBird band 4 (IR) as the green channel, and QuickBird band 2 (green) as the blue channel is given. This composite highlights changes in the infrared signature, primarily attributable to vegetation changes, between times of collection of the QuickBird and Duncan data. Figure 4.75b gives the same data with several linear features highlighted in yellow.



a **b**

Figure 4.75. Composite imagery at Kasita Town. a) Image composed of Duncan band 3 (IR) as red, QuickBird band 4 (IR) as green, and QuickBird band 2 (green) as blue. Interpretations are given in yellow. The project study area is outlined in white.

4.9.7. Lessons Learned

4.9.7.1. Data acquisition

Acquiring airborne data using fixed-wing aircraft over the study areas proved to be a challenge. The optimum time for data acquisition is in the spring when the ground cover is clear and the leaves have not emerged from the trees (*leaf-off* conditions). The contractor responsible for the data acquisition, GEOTEK, is located at Bay St. Louis, Mississippi. The distance from there to Ft. Riley, Kansas, Fort Benning, Georgia, and Silver Bluff, South Carolina proved to be difficult as rainy spring weather conditions often interrupted the data acquisition schedule. Sometimes the study area was cloudy or rainy. Other times the study area was clear but weather conditions in Bay St. Louis prevented take off. Most frustrating was when the study area was clear, the airport in Bay St. Louis was clear, but a severe band of storms between the two locations prevented data acquisition since the pilot could not navigate around the storms. The best approach for future acquisition missions would be to make use of local military aircraft, local contractors, or to acquire the data from a powered parachute (Section 4.8.1).

4.9.7.2. Time of acquisition of thermal data

Although it has been documented that, in general, the optimum time for thermal data acquisition is at thermal maximum and minimum (solar noon and pre-dawn) (Sever 1983, 1991; Luvall et al. 2005; Gonzalez 2005), it is recommended that before thermal data is acquired from the air, ground truth measurements be made to determine the optimum time for the data acquisition. All materials emit energy at different rates and are affected by temperature, pressure, and humidity. By comparing the emissions of known archeological features (such as ditches, hearths, kilns, and refuse areas) with the adjacent ground cover throughout the course of a 24 hour period, the optimum time to detect these features could be determined.

4.9.7.3. Mode of acquisition of thermal data

In large part due to the physical characteristics of the PalmIR 250, collection of clean, consistently usable data sets from an airborne platform proved to be problematic. Thermal data were recorded in the form of an image sequence by feeding the sensor signal to a frame grabber operating at 15 frames per second. The thermal instrument has both a relatively slow scene integration time and a relatively high sensor latency. These characteristics combine to cause a smearing artifact in the individual data frames. The severity of the artifact depends to a large extent on the aircraft's altitude and ground speed—slower speeds and higher altitudes lessening the effect—and, to a smaller extent, the level of thermal contrast through the target scene.

In light of this situation, the PalmIR 250 would probably be better utilized on a fixed boom, a tethered balloon-borne platform, or a potentially slower moving aircraft such as a powered parachute or helicopter (assuming a stable, vibration-free mount could be devised).

4.9.7.4. Site preparation

The grass was mowed at the study areas in order to allow better conditions for the ground-based geophysical surveys. It was also thought that this condition would be optimal for detecting archeological features through emerging springtime vegetation. This phenomenon, known as vegetation outcropping, has been used to detect archeological features throughout the world (Sever 2000, 1990; Deuel 1969; Crawford and Keiller 1928). This approach was successful in locating the 1910 hangar of the Wright Brothers at Wright Patterson Air Force Base (Sever 1998). In our analysis of the aerial remote sensing imagery for this study, however, we could see the effects of the mowing process and have not yet concluded if this improves or confuses the signal response to the aerial platform. A future investigation, that compares imagery over an area before mowing and after mowing, would resolve this issue.

5. RESULTS AND ACCOMPLISHMENTS

5.1. PROJECT HISTORY & MILESTONES

Deborah Harmon, Kenneth L. Kvamme, Fred Limp, University of Arkansas

5.1.1 Objectives

This project was initiated in December 2003 and the final report submitted May 2006. A detailed time line and task list is provided below. The projects objectives were all met and the specifics provided in summary form in this chapter and in detail in the body of the report. There were no go/no go decision points in the project.

A principal goal of this SERDP research project is a determination of remote sensing methods and techniques that work well individually, and that complement each other collectively when integrated, for the identification and evaluation of buried archaeological remains. Numerous methods for data integration are explored. With the exception of certain multi-band visualization techniques and overlays of vectors representing interpreted anomalies, most of the methods investigated have not been applied previously in archaeology. Several advanced computer graphic methods are explored, discrete methods that range from *Boolean* overlays to sums of categorized portrayals of sensor outputs and cluster analyses, continuous methods that include sums, products, ratios, principal components, regressions, and probability surfaces, intelligent knowledge-based systems like *C5.0* and *Cubist*, and Expert Systems approaches. This research demonstrates that certain integrating methods yield more information about the subsurface than others, but what may be realized in each approach may depend on overarching purpose. For example, some fusion techniques yield visually pleasing results that appear to well-combine available information, while others may seem less revealing but offer greater interpretive or predictive potential. In this process, the nature of similarities, differences, redundancies, and performance characteristics of results are examined. An important aspect of this research is an assessment of the added value of the fused product compared to traditional, individual-sensor based analysis.

The process of archaeological remote sensing as carried out in this project is a multi-step undertaking. The first stage, designed to meet the primary data integration goal, includes (1) remote sensing data collection (Section 4.2), (2) processing to clarify anomalies in individual sensors established within GIS databases (Section 4.4), (3) data fusion to integrate information from all sensors (Sections 5.2-5.11), (4) the definition of potentially "significant" cultural anomalies (Section 5.12), and (5) classifying the anomalies into likely types of cultural features (Section 5.12). The last might include indications of material composition and depth data. This is normally the end result of most remote sensing projects, and where archaeological fieldwork takes over. This project endeavored to go several steps farther with a *second* principal goal.

The second goal of this research project was very different and designed to meet an important criterion in the SERDP Statement of Need. Specifically, "ground-truth" testing was called for to demonstrate the feasibility of the proposed data integration technologies and associated procedures. Three additional project tasks were therefore designed. They include (6) a sampling design (Section 5.14) that allows (7) archaeological excavation of representative anomalies of each defined type to provide validation of remote sensing predictions about the subsurface (Section 5.16). This

validation phase often turned into a learning process, however, because the soils, geology and the archaeology in each site are unique, idiosyncratic, and confound predictability. In other words, remote sensing predictions cannot be perfect and a look into the ground through excavation offers additional insights that allow modifications to original predictions. Consequently, a final stage was defined that includes (8) modification of original remote sensing predictions, based on excavation findings (Section 5.17). Preparation of this report (9) was, of course, the ultimate task.

These operations were undertaken at each of four prehistoric and historic archaeological sites distributed across time and space in a wide diversity of environmental settings from South Carolina to New Mexico (Section 3.3). This allowed a means to ascertain the value of the methods investigated at very different archaeological sites with very different remains in very different environments. In so doing a better understanding of which methods consistently worked and offered useful results could be achieved, but this knowledge was also augmented by the considerable experience of SERDP Project team members.

This multi-year project was divided into seasonal and yearly tasks, with preliminary results, conference papers, and annual reports a constant part of the cycle. Owing to budgetary delays some project tasks were deferred. The project tasks and events accomplished in this project, or "milestones," are temporally organized in Table 5.1.

5.1.2 Objectives Summary

1. Acquire geophysical data, 4 sites – accomplished
2. Acquire aerial data, 4 sites – technically aerial data were acquired at all 4 sites, BUT anomalies of sufficient detail not generally found (except at Army City with thermal).
3. Acquire Satellite Data – successfully acquired, BUT anomalies of sufficient detail not found
4. Preprocessing of data – successful
5. Establish GIS databases – successful
6. Data integration/fusion – successful
7. Fieldwork for validation - successful
8. Evaluation of field results & remote sensing findings – successful

Table 5.1 *Subtask Schedule and Final Report Cross-Reference*

Task	Subtask	Completion Date	Report Discussion
1) Project initiation		11/2003	3.1, 3.2, 3.3
2)Aerial & satellite data acquisition			
	Acquire existing aerial data	03/2003	4.9
	Acquire satellite data	07/2003	4.9
	Acquire thermal data at Ft. Riley	04/2004	4.8
	Acquire thermal data at Ft. Bliss	06/2003	4.8, 4.9
	Acquire thermal data at Ft. Benning	06/2005	4.9
	Acquire thermal data at Savannah River	05/2005	4.9
3) Geophysical fieldwork			
	Acquire geophysical data at Ft. Riley	07/2002	4.6
	Acquire geophysical data at Ft. Bliss	10/2003	4.6
	Acquire geophysical data at Ft. Benning	01/2003	4.6
	Acquire geophysical data at Savannah River	01/2003	4.6
4) Preprocessing of data sets			
	Preprocess satellite imagery	10/2003	4.4
	Preprocess Ft. Riley geophysical data	01/2003	4.4
	Preprocess ATLAS data from Ft. Bliss	05/2003	
	Preprocess Ft. Bliss geophysical data	03/2004	4.4
	Preprocess Ft. Benning geophysical data	09/2003	4.4
	Preprocess Savannah River geophysical data	10/2003	4.4
	Complete preprocessing of thermal data sets from Bliss, Benning and Savannah River	09/2004	
	Preprocess thermal data from Ft. Riley	06/2004	
5) Establish GIS databases			
	Establish Whistling Elk and Mt. Comfort data bases	10/2002	
	Establish Ft. Riley GIS data base	10/2003	
	Establish Ft. Benning GIS data base	10/2003	
	Establish Savannah River GIS data base	10/2003	

	Prep Ft. Bliss geophysics and satellite data base	12/2003	
6) Data analysis and fusion/develop field testing program			
	Perform image segmentation and object classification analysis and fusion studies using existing Whistling Elk data	12/2003	4.5, 4.6, 5.3-5.13
	Perform initial image segmentation and object classification analysis on Ft. Riley data sets	12/2004	4.5, 4.6, 5.3-5.13
	Perform multivariate/statistical fusion on Ft. Riley data sets	08/2004	4.5, 4.6, 5.3-5.13
	Perform multivariate/statistical fusion on Ft. Benning data sets	11/2004	4.5, 4.6, 5.3-5.13
	Perform multivariate/statistical fusion on Savannah River data sets	12/2004	4.5, 4.6, 5.3-5.13
	Prepare predictive map for archeological studies at Ft. Benning	01/2005	4.5, 4.6, 5.3-5.13
	Prepare predictive map for Ft. Bliss	09/2004	4.5, 4.6, 5.3-5.13
	Prepare predictive map for Savannah River	01/2005	4.5, 4.6, 5.3-5.13
	Evaluate/compare predictions and field results at Ft. Benning	06/2005	4.5, 4.6, 5.3-5.13
	Evaluate/compare predictions and field results at Savannah River	06/2005	4.5, 4.6, 5.3-5.13
	Evaluate/compare predictions and field results at Ft. Bliss	06/2005	4.5, 4.6, 5.3-5.13
7) Technology transfer		06/2006	
8) Archaeology field testing / validation			
	Archaeological field work at Ft. Bliss	03/2005	5.13-5.17
	Archeological report and analysis on Ft. Bliss	09/2005	5.13-5.17
	Archaeological field work at Ft. Benning	03/2005	5.13-5.17
	Archaeological report and analysis on Ft. Benning	08/2005	5.13-5.17
	Archaeological investigations at Savannah River	03/2005	5.13-5.17
	Archaeological report and analysis at Savannah River	07/2005	5.13-5.17

9) Conyers consulting		12/2003	
10) Project analysis, evaluation, reporting			6
	FY 2002 Annual Report	12/2002	
	Presentation at SERDP Symposium	12/2002	
	FY 2003 Annual Report	12/2003	
	Presentation at SERDP Symposium	12/2003	
	FY 2004 Annual Report	12/2004	
	Presentation at annual SERDP meeting	12/2004	
	SERDP Final Report	6/2006	
	Submit Completed Fact Sheet to SERDP	6/2006	

5.1.3 Deliverable Products

The proposal *New Approaches to the Use and Integration of Multi-Sensor Remote Sensing for Historic Resource Identification and Evaluation* named 20 project deliverables. Nearly all are presented in the following pages, and some occur in previous reports, or in SERDP annual symposia and meetings. Section numbers of this report indicate where relevant materials may be found.

1. Description of ground-based field methods and instrumentation (Sections 4.2-4.3).
2. Description of satellite methods and instrumentation (Section 4.7).
3. Description of data processing methods (Section 4.4).
4. Graphical and data description of remote sensing findings from each study site (Section 4.6).
5. Interpretation of remote sensing findings from each study site (Section 4.6).
6. Description of data integration/fusion methods (Section 5.2-5.11).
7. Data fusion imagery from each study site (Section 5.2-5.11).
8. Operational data fusion software system. (ESTCP Project SI-0611)
9. Interpretation of integrated remote sensing findings from each study site (Section 5.12).
10. Description of archaeological field validation research design (Section 5.14).
11. Description of archaeological field validation methods (Section 5.15).
12. Description of archaeological field validation results and multidimensional remote sensing accuracy assessments (Section 5.16).
13. Discussion of kinds of archaeological features found and missed, false positives and negatives (Section 5.16).
14. Discussion of benefits of multidimensional remote sensing data integration (Section 6).
15. Discussion of data redundancies and cost-benefits within each ground, air, or space domain and between domains (Section 6).
16. Recommendations of the best subset of instrumentation necessary for future work of this kind (Section 6).

17. Discussion of environmental characteristics and their effect on detection (Section 6).
18. Annual technical reports. (Submitted to SERDP, available online at http://www.cast.uark.edu/cast/serdp/index.html)
19. Final technical report. (Submitted to SERDP, available online at http://www.cast.uark.edu/cast/serdp/index.html)
20. Presentations or poster at annual SERDP symposia and meetings. (Appendix B Technical Publications)

5.2. PREPARING DATA FOR INTEGRATION

Kenneth L. Kvamme, University of Arkansas

5.2.1. Site Selection and Methods

Although several project sites yielded excellent geophysical responses, the site that receives principal focus in an examination of data integration methods is Army City. By utilizing a single site as a point of reference, comparisons between methods will be facilitated, greatly reducing confusion that might arise by inspecting the nuances of so many methods across a large number of sites (data fusion results for other sites are presented in Section 5.11). At Army City, geophysical results were good to excellent in *six* geophysical dimensions—magnetic gradiometry, electrical resistivity, soil conductivity, magnetic susceptibility, ground-penetrating radar, aerial thermography—more than at any other project site (see Section 4.6). Moreover, Army City offers the only useful aerial data set in the entire project, allowing exposition of integrations based on ground *plus* aerial data. At Army City a relatively large area was surveyed by all methods—1.6 ha—addressing a fundamental remote sensing requirement: survey of large areas that is necessary to facilitate recognition of broadly distributed cultural patterns (Section 4.1.3). The spatial resolution at Army City was high, allowing excellent detail and detection of small features. Finally, a number of historic photographs exist that greatly aid interpretation, as does an historic plat map showing the locus (and names) of historic streets (see Section 3.3.4.1). Taken together, Army City represents an ideal site for examining methods of integration. It possesses high geophysical dimensionality, uses ground and aerial data, offers large-area coverage, high spatial resolution, and relative ease of interpretation owing to historic photos and maps and because cultural anomalies are more easily recognized in this site with features similar to those in our own society.

In the Sections 5.3-5.10 a host of very different methods are investigated for integrating multi-sensor data at Army City. A subsequent Section 5.11 examines a more limited number of data integrations at other sites. Methods of data integration span a wide variety of approaches and computational domains. Depending on the approach used, different goals or benefits may be realized. Many methods offer improved ways to visualize the totality of subsurface evidence by portraying multiple geophysical dimensions simultaneously. Many help delineate anomalies with greater clarity. Several help to improve theoretical understandings of relationships between the various geophysical methods and archaeological deposits through graphical and statistical means. In many cases, the identification of what kind of archaeological feature an anomaly represents is improved, which aids the interpretive process. Some techniques may also augment prospecting by making new anomalies not visible in any single data set clear.

Some methods are relatively "simple" to understand and implement; others are extremely complex algorithmically. To impose some form of order, these methods are grouped into *seven* basic integrating domains and ordered from the simple to the complex. The seven domains include:

1. *traditional interpretive approaches* for integration that utilize simple graphic overlays of anomaly vectorizations (Section 5.3),
2. *graphic methods* that range from overlays to color composites and translucent data layers (Section 5.4),
3. *discrete methods* that utilize diverse approaches from Boolean methods to cluster analyses (Section 5.5),
4. *simple mathematical operations* including sums or ratios of data sets (Section 5.6),
5. *multivariate methods* from principal components to regression techniques (Section 5.7),
6. *intelligent knowledge-based algorithms* that employ advanced machine-learning algorithms, context (Section 5.8), and
7. *expert systems* based on an expert's knowledge of geophysical responses and Army City's archaeology (Section 5.9).

5.2.2. Pre-processing for Data Integration

The considerable initial data pre-processing of Army City's geophysical data was discussed in Section 4.4. In order to undertake data fusions between the very different data sets, several additional pre-processing steps were necessary. The magnetic gradiometry and susceptibility surveys exhibited large dipolar values (dipolar anomalies that also occurred in the EM conductivity data were less extreme and therefore ignored). Each joint of the many buried pipes revealed by magnetic gradiometry exhibited adjacent positive and negative poles of extreme value that represented the same feature, which would negatively influence statistical or visual relationships in data integrations. Although an advanced technique known as "reduction to the pole" may theoretically be employed for reducing dipolar anomalies to source points (Telford et al., 1990:109), software reconstructions of such linear objects as pipes are difficult to achieve, so absolute values of these data were taken instead.

The six geophysical data sets were field-sampled at different spatial resolutions ranging from .025 x .5 m to .5 x 1 m for the ground-based surveys to about .1 x .1 m for the aerial thermography (Section 4.8). To facilitate data integration, all data sets were resampled to a common spatial resolution of .5 x .5 m (the commonest resolution) through bilinear resampling in order that common boundaries would match in the predominantly pixel-based approaches. Additionally, different measurement scales, data ranges, and even distributional forms are exhibited by the geophysical data. In order to side-step potential difficulties that such a mixed bag might introduce to data integrations (e.g., one variable might dominate because of larger measurement values or exert undo influence due to outliers), the measurements were subjected to a series of transformations. Power or logarithmic functions were applied to approximately normalize each distribution and the data were standardized with $z = (x_i - \mu)/\sigma$ (where μ is the mean, σ the standard deviation), yielding $\mu=0$, $\sigma=1$, and data that typically range between +/- 4σ. This process is

illustrated for the Army City resistivity data in Figure 5.1. Further pre-processing required by specific data fusion techniques is discussed in sections below.

The six pre-processed data sets are illustrated in Figure 5.2 with principal cultural features discussed in Section 4.6.1 labeled. These data represent the start-point to data integrations that proceed from the simple to the complex, beginning with manual interpretations and computer graphic techniques, and moving through mathematical transformations to intricate statistical and intelligent knowledge-based manipulations.

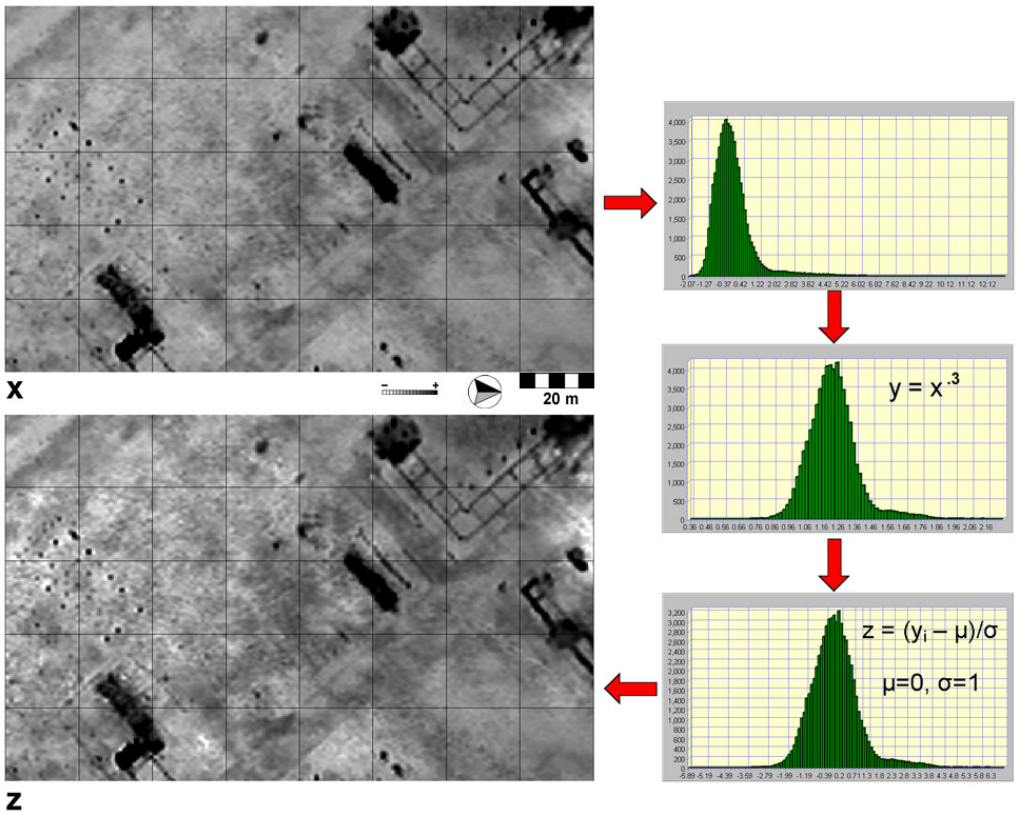


Figure 5.1. *Normalizing transform and standardization of the Army City resistivity data. x) Initial pre-processed resistivity data. z) Data after transformation. Both data sets are displayed by clipping at +/- 3 s.d.*

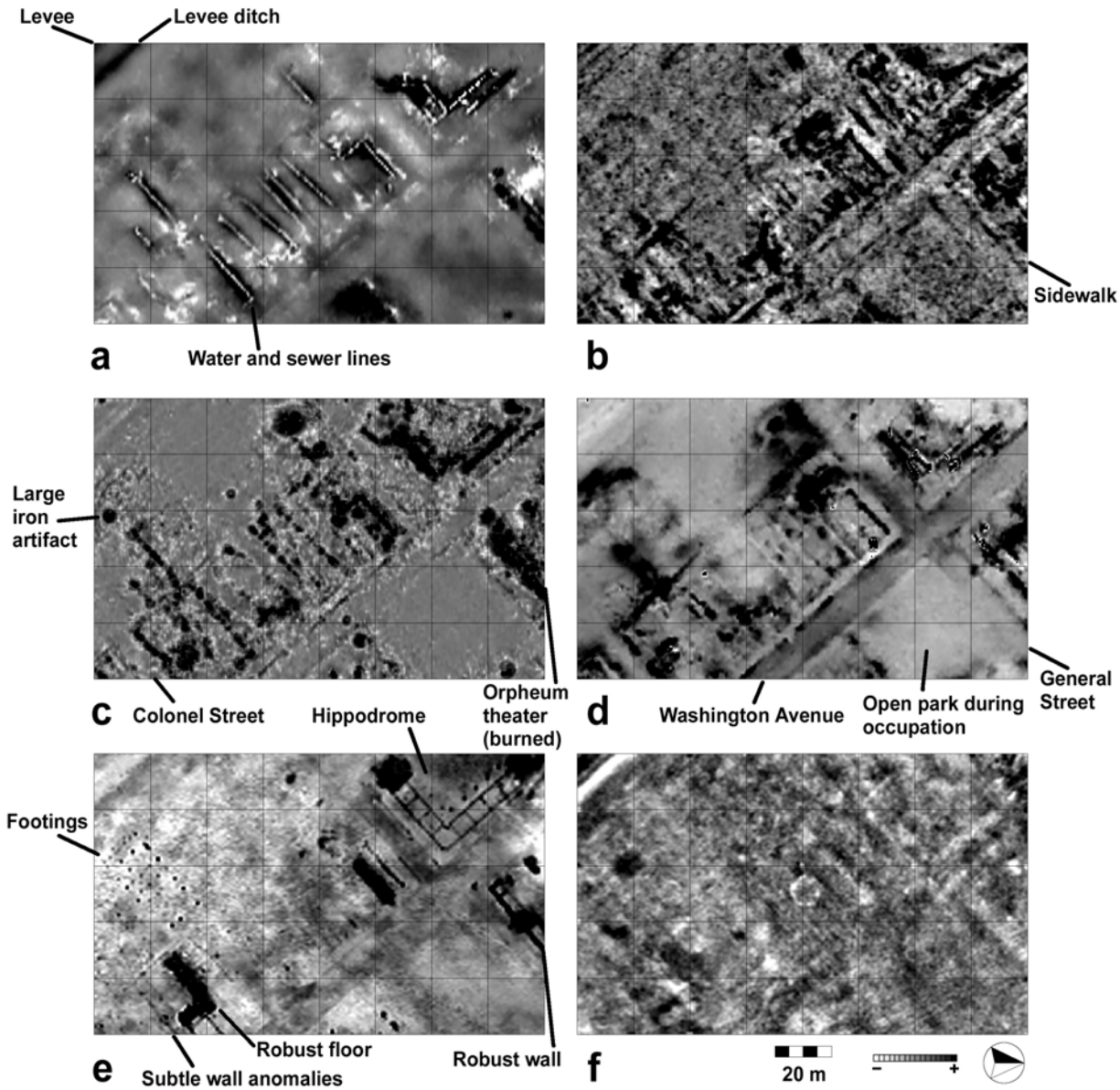


Figure 5.2. *The six pre-processed geophysical data sets from Army City: a) EM conductivity, b) GPR time-slice (20-40 cm below surface), c) absolute magnetic gradiometry, d) absolute magnetic susceptibility, e) electrical resistivity, f) thermal infrared.*

5.3. TRADITIONAL INTERPRETIVE APPROACHES FOR DATA INTEGRATION

Kenneth L. Kvamme, University of Arkansas

Geophysical findings are traditionally obtained through subjective interpretations of the data combined with deductive reasoning, whether the data are uni- or multi-dimensional. Successful interpretations of likely cultural features rely on expertise in the local archaeology and knowledge of corresponding archaeological signatures in geophysical data. The methodology rests largely on visual interpretation of geophysical

maps followed by the laborious process of *vectorization*—the manual tracing of likely cultural anomalies on paper or a computer screen using GIS or equivalent digitizing software (Burrough and McDonnell 1998:88). This subjective procedure is extremely time-consuming, and can be viewed as the conversion of continuous geophysical raster data to discrete point, line, or area vectors (in GIS parlance) that represent significant cultural features. In other words, the end product is a series of interpreted maps depicting the locations of likely cultural features. If multiple geophysical surveys have been conducted then a separate map is produced showing interpreted cultural anomalies in each one. If combined and overlaid, a form of data integration is represented. Many examples of this practice exist in the literature (Buteux et al. 2000:77; Gaffney et al. 2000; Neubauer 2004:163) and the approach has justifiably been presented as a goal of archaeo-geophysical surveys (see Schmidt 2001:23).

At Army City, each geophysical data set of Figure 5.2 was carefully examined to define potential cultural anomalies based on their form, distribution, limited historical maps, photographs (Figure 3.2), and excavation results. Anomalies interpreted as culturally significant were then vectorized data set-by-data set through heads-up digitizing (Figure 5.3a-f). This process generated discrete line and polygon vectors, each in two classes: robust or more secure anomalies and those more subtly indicated (they are distinguished in Figure 5.3). GIS methods easily allow vector data to be overlaid. In so doing, a form of data integration is achieved, because all interpreted information from all sources is simultaneously portrayed (Figure 5.3g). This integration represents a point of departure from methods that follow, however. In contrast to this manually intensive and subjectively driven technique, subsequent approaches present, for the most part, *algorithmic* and more automated means for data integration.

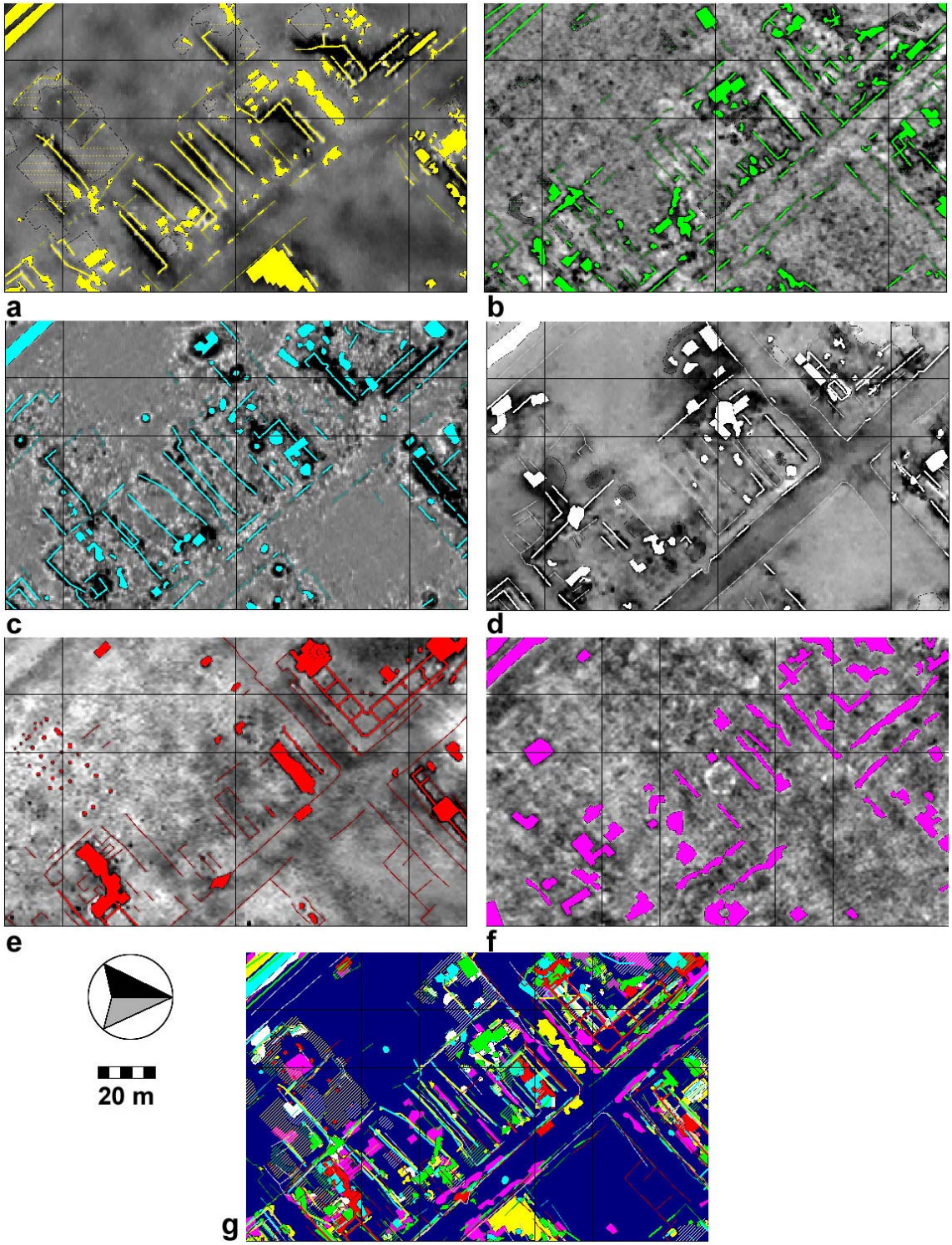


Figure 5.3. *Interpretive vectors overlaid on (a-f) geophysical data and (g) each other. a) Conductivity, b) GPR, c) magnetic gradiometry, d) magnetic susceptibility, e) electrical resistivity, f) thermal infrared, g) composite overlay integrating the data. KEY: yellow=conductivity; green-GPR; cyan=magnetic gradiometry; white=magnetic susceptibility; red=resistivity; magenta=thermal infrared; robust anomalies=solid bright colors; subtle anomalies=cross-hatching or dull-colored line vectors.*

5.4. GRAPHICAL METHODS FOR DATA INTEGRATION

Kenneth L. Kvamme, University of Arkansas

Several computer methods of graphic portrayal have been applied for decades that yield effective data fusions. In fact, some are so common they may not be understood as fusions. While many display variations exist, only four principal ones are examined here.

5.4.1. Graphic Overlays

Four methods are most commonly employed for displaying continuous map data: isoline contours representing measurements of equal value, gray or color scale variations signifying measurement magnitudes, pseudo-three-dimensional views where the vertical axis represents measurement size, and shadowing to indicate the steepness of gradients in three-dimensional surfaces. Each of these methods may be used to represent a single data set, but several data sets may be simultaneously portrayed through overlaying, offering a means of data fusion. An effective combination utilizes one data source as a color or gray-scaled background image, with a second one overlaid with isoline contours. For instance, in Figure 5.4a the background electrical resistivity data tend to highlight Army City's buried concrete foundations, building footings, and other resistant features, while the overlaid EM conductivity data locates water or sewage pipelines associated with these structures (compare Figures 5.2a,e). A second useful display is illustrated in Figure 5.4b where magnetic susceptibility is portrayed as a three-dimensional surface that mostly reveals streets and some structural evidence (with variations emphasized through shading in gray), and GPR is overlaid in color that generally indicates robust walls and floors (compare Figures 5.2b,d). These data fusions instantly convey profound similarities and differences in anomaly distributions and emphasizes that various methods measure very different properties of the ground.

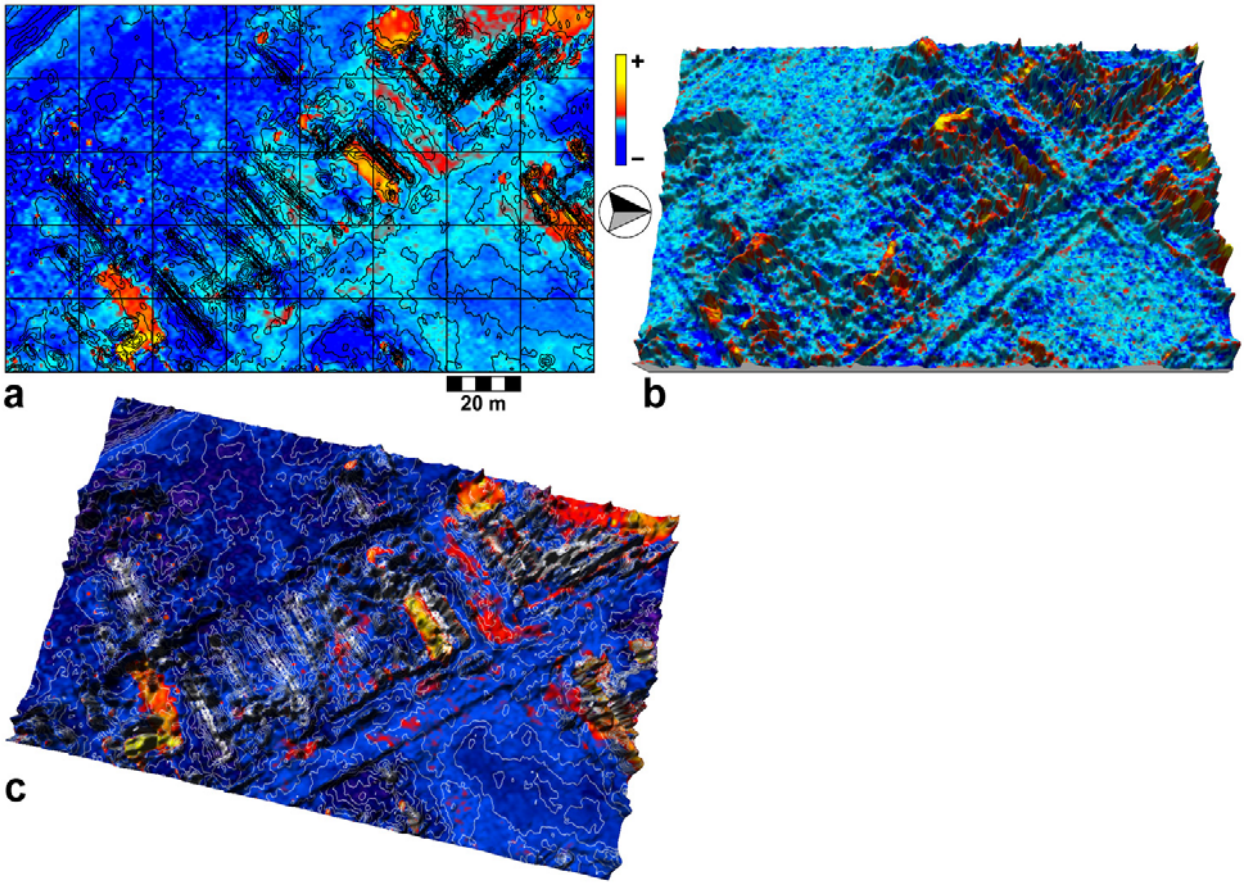


Figure 5.4. *Graphic overlays as data integration. a) Conductivity contours (black) overlaid on resistivity as a color raster. b) GPR, in color, overlaid on a shadowed magnetic susceptibility surface. c) Conductivity contours (white) overlaid on resistivity as a color raster on top of shadowed magnetic susceptibility surface.*

Graphic overlays can obviously be carried to portrayals of more extreme dimensionality. For example, a surface might represent one dimension, shadowing a second, a color overlay a third, and contouring a fourth. Such maps rapidly become too complex for ready understanding, however, with 2-3 dimensions probably offering a practical limit to what the human brain can effectively process. Yet, with care, strides might be taken in that direction. Figure 5.4c illustrates a carefully constructed, three-dimensional graphic where soil conductivity contours, in white, are overlaid on a colorized resistivity raster, which in turn rests over a shadowed magnetic susceptibility surface. If one keeps in mind that all black shadows point to susceptibility, that all white lines indicate conductivity, and that all color represents resistivity, simultaneous consideration of all three data sets may be made.

5.4.2. RGB Color Composites

Red-green-blue (RGB) color composites have been a standard form of display of satellite imagery for decades (e.g., see Schowengerdt 1997). In this visualization method, each of three data sets is assigned a primary color and the images are then combined to create color mixes that span the full visual spectrum. When more than three dimensions

are available many primary color assignments become possible, and each can yield different insights. For instance, with six Army City data sets (Figure 5.2) there are $C_{6,3}$=20 possible groups of three, with 3!=6 permutations of red, green, and blue for each, yielding 120 possible displays! One such RGB composite is illustrated in Figure 5.5c that well shows the site's commercial core and relationships between electrical resistivity (red), conductivity (green), and magnetic susceptibility (blue).[2] Three other color composites formed from the same data sets, but with changed color assignments, are given in Figure 5.5e-g for comparison purposes.

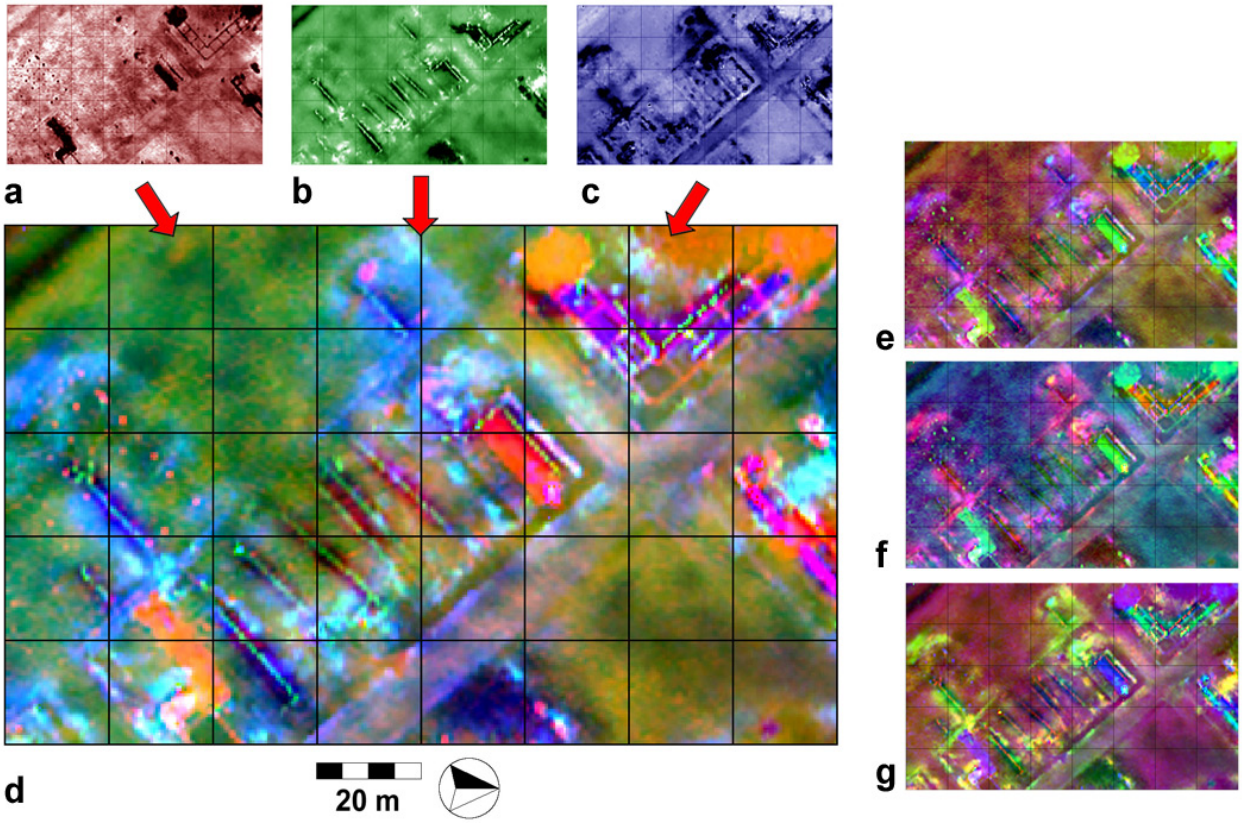


Figure 5.5. *RGB color compositing. a) Resistivity in red, b) conductivity in green, and c) magnetic gradiometry in blue to create d) an RGB composite. e-g) represent three of the five alternative RGB composites using the three inputs* a-c.

Knowledge of color theory and the RGB color model substantially enhances interpretation. Bright red, green, or blue in a color composite points to high values of the variable assigned to the respective color, but low values of the other variables. White indicates large measurements in all three data sets, while black points to the reverse. Yellow means high values in the variables associated with green and red, magenta to elevated measurements in red and blue, and cyan to large values in green and blue. Consequently, at least eight simple interpretations are possible, plus numerous shades of gray and a myriad of colors between. In Figure 5.5c the Hippodrome, other foundations, and building footings are exclusively revealed by electrical resistivity (red), streets are indicated primarily by magnetic susceptibility (blue), and pipelines by EM conductivity (green); at the same time, certain walls and floors are robustly positive in the three

primaries (making white) while a recent drainage ditch in the southwest corner (upper left) shows a uniform negative result (creating black). The RGB color composite is visited again in later sections to fuse input images of greater dimensionality.

5.4.3. Translucent Overlays

Recent advances in computer graphics allow color overlays of many more dimensions than the three of the RGB model through use of translucencies, but the approach lacks similar theoretical grounding.

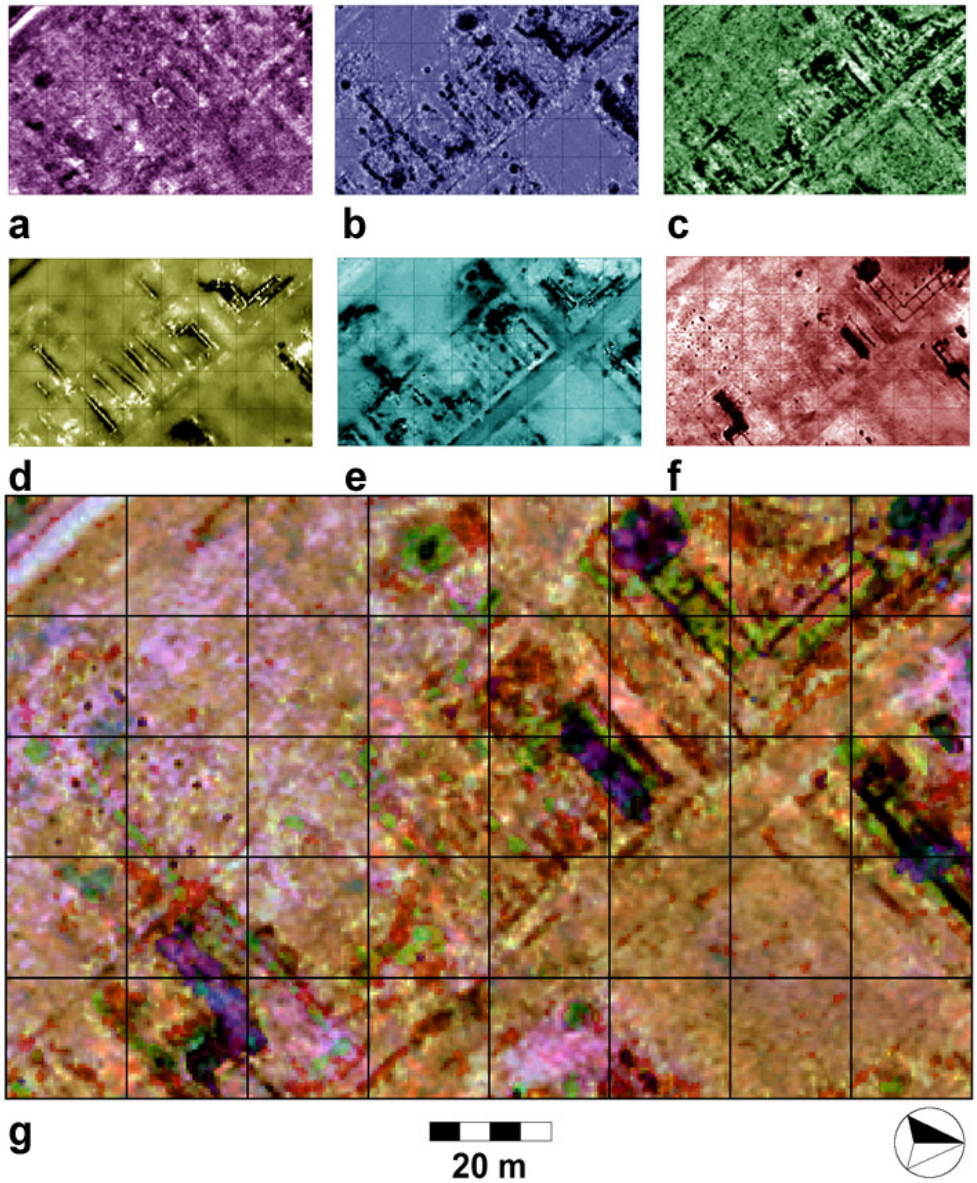


Figure 5.6. *Translucent overlays. a) Thermal infrared, b) magnetic gradiometry, c) GPR, d) conductivity, e) magnetic susceptibility, f) resistivity, g) composite of* a-f *with* a *as an opaque base, and the remaining images as 40-60% translucencies.*

A very different color is assigned to each of *k* data sets; *k*-1 is then made translucent and overlaid on a *k*th opaque image beneath. Using this methodology, all six data sets of Figure 5.2 are simultaneously overlaid in Figure 5.6g, resulting in a more comprehensive depiction of anomalies than the three-dimensional RGB composites of Figure 5.5. A general limitation is that high *k* tends to produce a "muddy" looking result, and layers lying beneath several others may be nearly obscured. In general, translucent overlays of 3-4 images may produce clearest results.

5.4.4. Evaluation

Computer graphic solutions for data integration are easy to implement and effectively combine information from several disparate sources in easy to interpret displays. They offer a ready means to generate complex visualizations of the subsurface. Their weakness lies in their relatively low dimensionality—typically only 2-3 dimensions may be represented (or perhaps somewhat more using translucent overlays). Moreover, these methods are purely descriptive, capable of combining only what is able to be maped in contributing data sets, and they yield only images, not new data that may subsequently be analyzed.

5.5. DISCRETE INTEGRATING METHODS

Kenneth L. Kvamme, University of Arkansas

Discrete methods, where data occur in distinct steps, levels, or categories, offer a number of advantages for data integration. Classes in categorical maps are mutually exclusive with well-defined boundaries that remove ambiguity about group membership and anomaly limits. The data are also easily combined and manipulated allowing use of relatively simple, yet powerful, methods available in common GIS and other software. Further pre-processing of the data is required in order for the methods to perform optimally.

5.5.1. Pre-processing: Generating Binary Data

Binary data lie behind a principal *modus operandi* in the realm of GIS modeling, forming the inputs to what is known as *Boolean* operations, a primary data-integrating tool (Burrough and McDonnell 1998:164-166). These approaches require the raw continuous geophysical data output by field instrumentation to be reclassified to a binary state. In the present context, a "1" may indicate the presence and "0" the absence of an anomaly. This binary representation is achieved by selecting appropriate thresholds near the tails of a statistical distribution and mapping only those locations with more extreme values as unity, zero otherwise (Section 4.4).

Thresholding of raw geophysical data might appear to be a straightforward operation, but problems arise in application. Broad regional trends arising from soil changes or underlying geology frequently occur that can mask cultural anomalies. Removing these trends, or "flattening" the data prior to binary classification, is therefore warranted. Although the Wallis filter (Scollar et al. 1990) is a popular trend-removing algorithm in archaeological remote sensing, we found that high-pass filters or subtraction of low-order polynomial trend-surfaces performed adequately. A second problem occurs in relatively "quiet" areas of the data where the signal-to-noise ratio is low. In these

contexts, low-pass filters can help to consolidate and strengthen weak anomalies, making them more cohesive and apparent. The effects of these operations are illustrated for the Army City resistivity data in Figure 5.7b. An appropriate threshold was then selected that visually appeared to best define cultural features, and a binary data set was generated, illustrated for the resistivity data in Figure 5.7c (compare this binary classification with the one derived from un-flattened data in Figure 4.25.1). These transformations were applied to each data set in Figure 5.2 and binary data sets were generated (Figure 5.8a-f). These results were then subjected to the following operations.

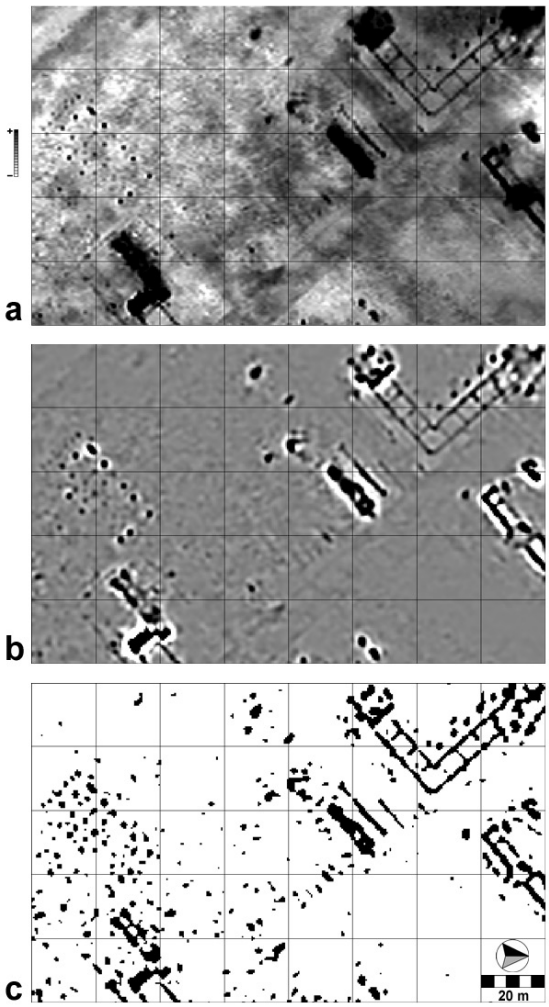


Figure 5.7. *Flattening and binary classification of the Army City resistivity data. a) Normalized and standardized data (from Figure 5.1), b) after flattening and consolidation with high- and low-pass filters, c) binary classification based on a +2 s.d. threshold.*

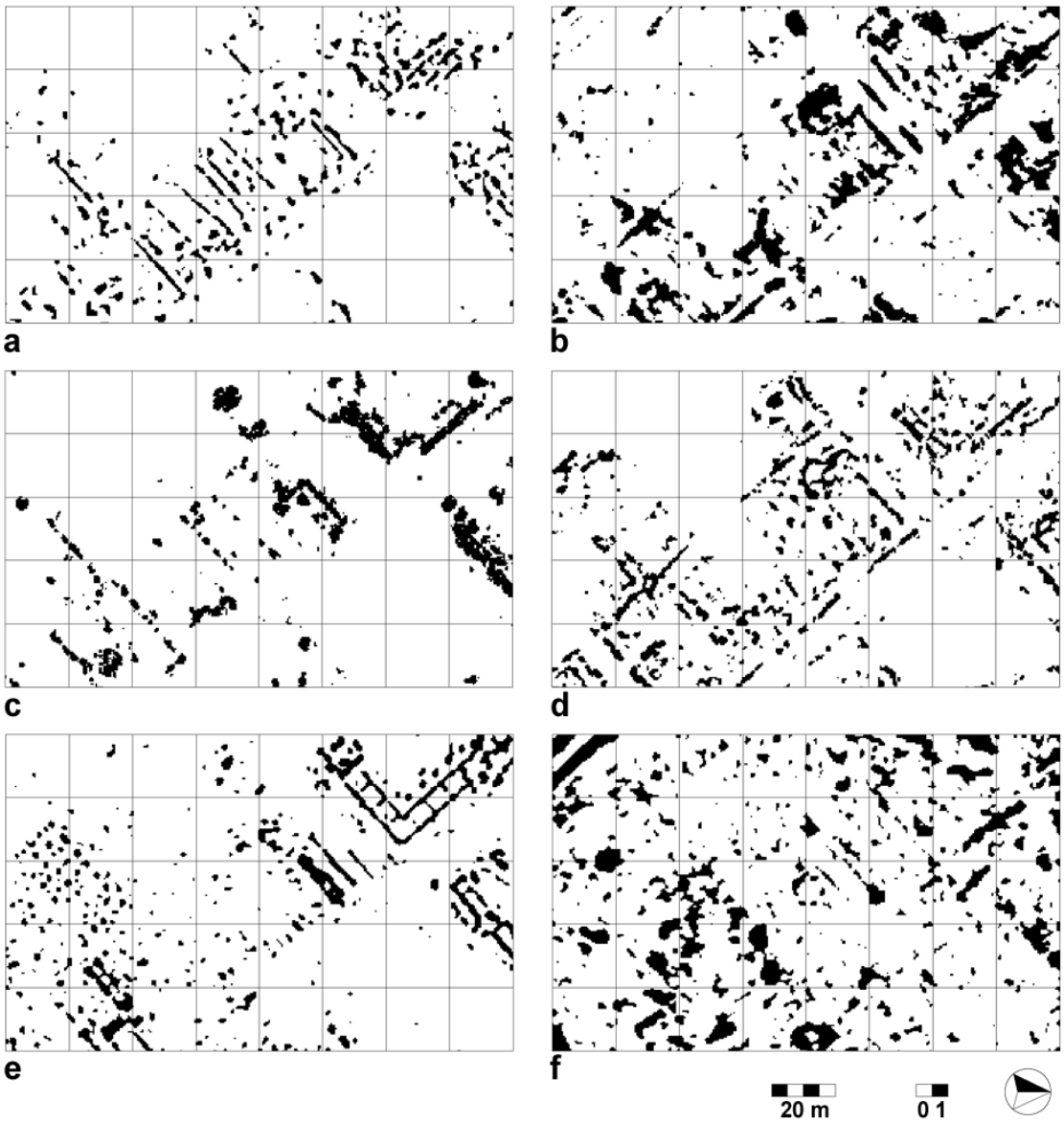


Figure 5.8. *Binary representations of principal geophysical anomalies at Army City: a) EM conductivity, b) GPR, c) magnetic gradiometry, d) magnetic susceptibility, e) electrical resistivity, f) thermal infrared.*

5.5.2. Boolean Union

A Boolean union occurs when *at least* one of the inputs is coded as "1." Its mapping shows all places where anomalies are indicated by one or more methods. In other words, the union simultaneously shows the loci of all major anomalies in all data sets (Figure 5.8a-f). As such, the outcome of this data integration produces a primary map of interest (Figure 5.9a). As a combination of all anomalies from all geophysical sources, the union can map a relatively large area: 39.7 percent of the region (.63 of 1.6 ha) in Figure 5.9a is anomalous by this method. This large area is in part due to the many prospecting methods forming the union, and the relative distinctiveness of each (see below). With fewer data sets in a union, smaller areas will generally be mapped.

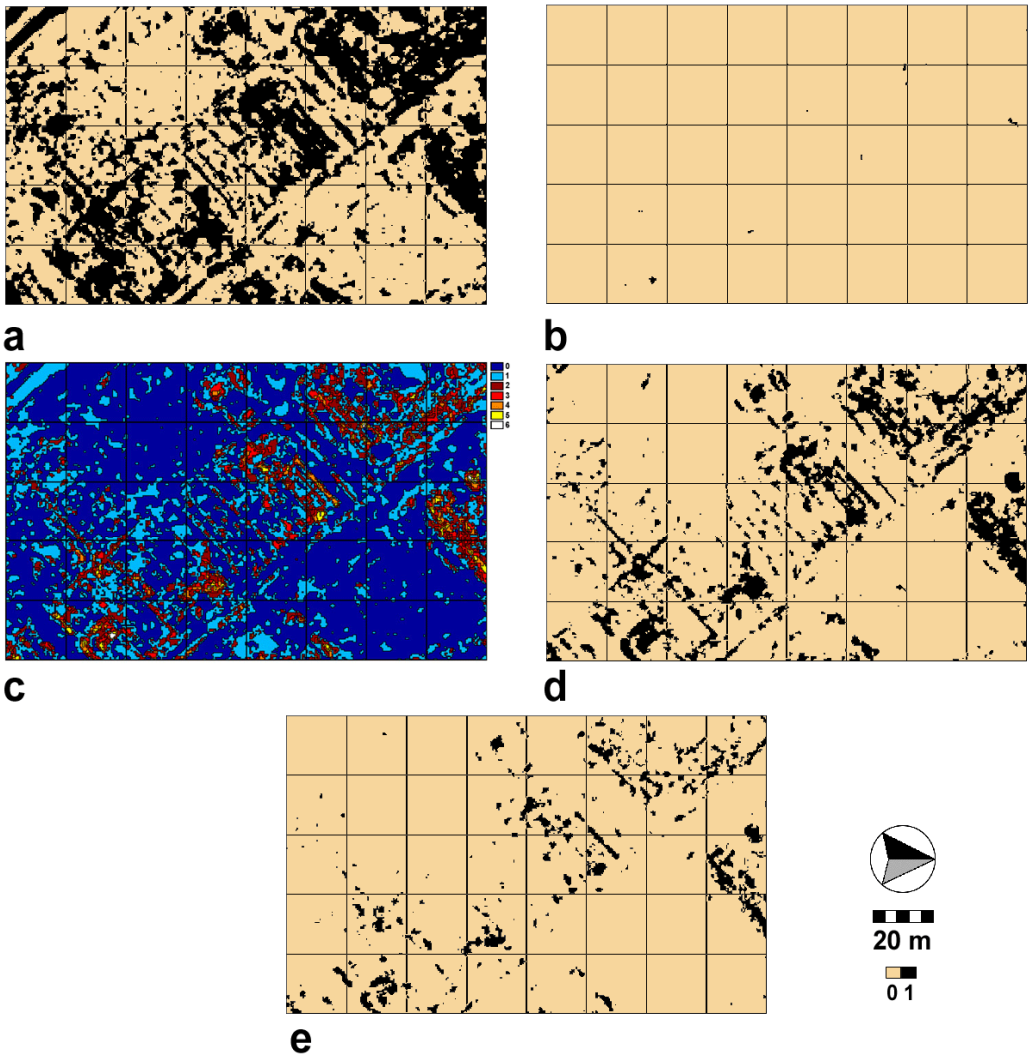


Figure 5.9. *Discrete data integrations at Army City. a) Boolean union, b) Boolean intersection, c) binary sum, d) anomalies indicated by at least 2 methods in* c*, e) anomalies indicated by at least 3 methods in* c.

5.5.3. Boolean Intersection

The Boolean intersection is also of interest because it reveals places where defined anomalies *simultaneously* occur in all geophysical data sets. In other words, for a result to be "true" a "1" must occur in *every* input data source. At Army City only 11.5 m^2 (.07 *percent* of the region's 16,000 m^2) meets this condition, representing targets with very different properties in all dimensions (Figure 5.9b). This restricted result is not surprising because subsurface entities that cause anomalies to be recorded will rarely yield strong contrasts concurrently in so many geophysical dimensions (magnetic, electrical, thermal). Studies using lower dimensionality (e.g., 2-3 methods) will typically yield less limited results.

5.5.4. Binary Sum

A *sum* of binary maps offers several distinct advantages over Boolean operations. With *k* inputs, the result can theoretically range between 0-*k*, yielding *k*+1 classes. Zero occurs at locations where all binary geophysical inputs are coded zero while *k* (the maximum) will result where all inputs are coded as unity. Consequently, imbedded within this method are the Boolean intersection and union because locations achieving *k* represent the former and all non-zero loci the latter. The sum also represents a *ranking* of anomalies revealed by zero of *k* methods, by one of *k* methods, and so on up to *k* of *k* methods. As such, it can be interpreted as an anomaly "confidence" map, where larger values signify that more geophysical methods reveal the same anomaly (Figure 5.9c). This map is therefore much richer in content and detail than the Boolean union or intersection (Figure 5.9a, b). It is noteworthy for its simplicity, and well indicates the layout of Army City's principal cultural features. It shows that many anomalies, particularly certain foundations, floors, pipelines, and gutters, are revealed by several geophysical methods, while other anomalies—certain floors, street gutters, walls—are indicated by only a single method.

5.5.5. Thresholds on Binary Sum

The binary summation (Figure 5.9c) can itself be reclassified to indicate, for example, all places where two, three, or more (to *k*-1) sensors reveal a major anomaly. This approach also yields binary outcomes, but ones that lie between the extremes of the Boolean union (one or more indications of anomalies) and intersection (*k* indications of anomalies). In other words, a threshold is selected in the binary summation map between 1 and *k* that may yield a more informative result than either Boolean outcome. To illustrate, mapping anomalies indicated by *two* or more geophysical methods (Figure 5.9d) well defines Army City's major features (foundations, floors, pipelines, street outlines) by eliminating less secure anomalies revealed by only a single method. The mapping of anomalies, indicated by three or more methods, shows particularly robust features of the site (Figure 5.9e).

5.5.6. Polychotomous Integrations

The foregoing discrete approaches, based on dichotomous classifications of the geophysical data, may be carried a step further by moving to polychotomous classifications of each geophysical data set. This offers an advantage because with a greater number of classes and class thresholds more subtle anomalies might be defined, in contrast to two-class applications where only the most robust anomalies tend to be delineated (Figure 5.8). The geophysical data might be classified into three classes (e.g., low, medium, high measurements), four classes (low, medium-low, medium-high, high), and so on, *ad infinitum*, up to the logical conclusion being the original continuous measurements represented in Figure 5.2. Owing to the underlying continuous nature of the geophysical data, the resulting classes are ordinal. Simple summations of the ranked classes might be made to yield a modeling outcome that is quite popular in GIS modeling circles, where it forms the basis of "multi-criteria decision modeling" (Burrough and McDonnell 1998). In the present circumstance, this approach can form effective data integrations.

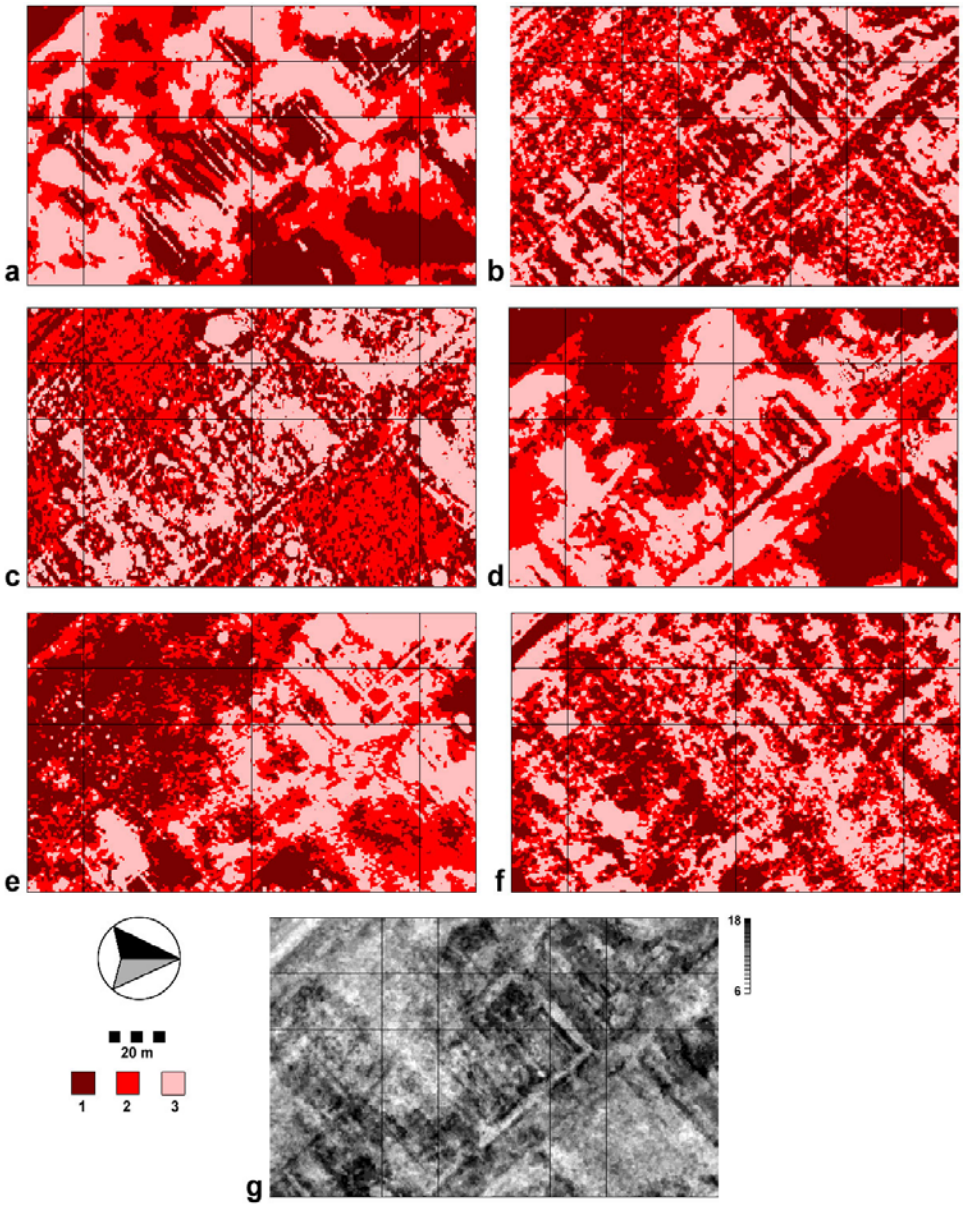


Figure 5.10. *Integration of ranked polychotomous geophysical data by summation. a) Conductivity (inverted scale), b) GPR, c) magnetic gradiometry, d) magnetic susceptibility, e) resistivity, f) thermal infrared, g) summation model ranging from 6-18 in magnitude.*

Realizing that infinite permutations are possible, only three classes are considered here for illustration purposes (a 5-class solution is illustrated in Section 5.11). In other words, each geophysical data set is classified into three ranked classes representing low, medium, and high measurements (and numbered 1-3), with each class of approximately equal area (i.e., the data distribution was divided into thirds; Figure 5.10a-f). Other class boundaries based on equal-intervals or standard deviations are, of course, possible. An integrated model is achieved by *summing* the six data sets with the outcome, in this case, ranging from 6 (the minimum) to 18 (the maximum), illustrated in Figure 5.10g. In general, such models appear very continuous to the eye (with the outcome in Figure

5.10g having 13 ranked classes), the more so as the number of classes and geophysical inputs is increased. In fact, the outcome achieved here closely compares with a summation model based on the continuous data of Figure 5.2 (with an "infinite" number of categories), described in a section below. In view of this result, it is argued that the summing of polychotomous classifications offers a poor second alternative to the summing and other manipulations of continuous data, described below.

5.5.7. Cluster Analysis Methods

Cluster analysis refers to a series of algorithms designed to define natural groupings in a body of multivariate data such that each one is more or less homogeneous and distinct from others (Davis 2002:487). In satellite remote sensing this method represents a form of "unsupervised classification" where an algorithm is data-driven to define intrinsic classes (Schowengerdt 1997). It is most ideally suited for application to continuous data, so the initial pre-processed data of Figure 5.2 are utilized. With a goal of defining natural classes, the result is discrete, however.

The *k*-means cluster analysis algorithm was selected because it operates rapidly on large data sets (here, n=64,000 in each of six dimensions). Beginning with a user-specified number of clusters, k, the algorithm places k arbitrarily located means in the six-dimensional measurement space, computes the Euclidean distance between each point and the nearest mean (hence the importance of commensurate measurement scales achieved through standardization), and computes a sum of squared error (SSE) statistic, the sum of squared distances from the respective means. At each step the algorithm iteratively moves the *k*-means about the measurement space until the SSE is minimized, which indicates that an optimal partitioning into k classes is achieved with each case assigned to the closest cluster mean. Of course, one generally has no idea about how many classes might truly exist, which poses a real problem in the present context. One might make an argument for only two classes ("anomaly" versus "background"), three classes ("positive anomaly," "negative anomaly," "background"), or many classes, perhaps hoping the algorithm can differentiate between such different anomaly types as "wall," "floor," "building footing," "pipeline," "gutter," "street," and "background," for example.

The standardized data were subjected to *k*-means cluster analyses, with k=2-8. It was found that with $k \geq 7$, small and insignificant clusters containing few pixels were defined, forming small micro-groups (Table 5.2). Building clusters from k=2 upwards yields excellent insights about the structure of the data, and facilitates overall interpretations. The k=2 solution illustrates the fundamental dichotomy of anomalous locations versus the natural background, as predicted (Figure 5.11a). The k=3 solution maps a background class, but splits the anomaly class of the k=2 result into two groups that appear to define positive floor, wall, and street gutter anomalies, versus "negative" pipeline anomalies revealed by EM conductivity and magnetic gradiometry (Figure 5.11b). The four-cluster solution adds more insights because at this level of partitioning the "background" of the previous two iterations is divided according to thermally indicated anomalies (Figure 5.11c; compare Figure 5.2f). This thermally divided background remains nearly constant under k=4-8 classes and is explained by analyses below that indicate the thermal infrared data to be the dimension most independent of the others.

The *k*=5 and *k*=6 solutions are very similar in that they begin to illustrate greater anomaly detail and define specific classes of archaeological features (Figure 5.11d,e). In some cases, it is difficult to assign specific interpretations to the defined classes, but comparison against the primary data of Figure 5.2 suggests several possibilities. Classes 1 and 2 (dark browns) represent the background, but divided according to thermally indicated positive and negative anomalies (compare Figure 5.2f). Class 3 (white) appears to represent robust and broad floors, best indicated by the resistivity data, but a number of smaller walls and footings also occur in this class (compare Figure 5.2e). Class 4 (yellow) generally points to robust linear pipe features indicated by conductivity and magnetic gradiometry (compare Figure 5.2a,c), while Class 5 (red) corresponds with broad areas of high magnetic susceptibility surrounding some buildings and along street gutters (compare Figure 5.2d), probably reflecting burned areas in the former and magnetic enhancement in the latter. Finally, Class 6 (orange) appears best correlated with robust GPR anomalies (compare Figure 5.2b), which appear to be associated with certain kinds of walls and likely floor areas. While these suggestions are insightful, such detailed interpretations must wait until Section 5.16 for validation by field excavation.

Table 5.2. *K-means cluster analysis statistics (n=64,000).*

Number of Clusters, *K*	**SSE**	**Smallest class size (pixels)**
2	309,911	13,691
3	277,117	5,717
4	250,072	5,262
5	233,364	4,270
6	217,557	3,259
7	223,566	21
8	208,201	39

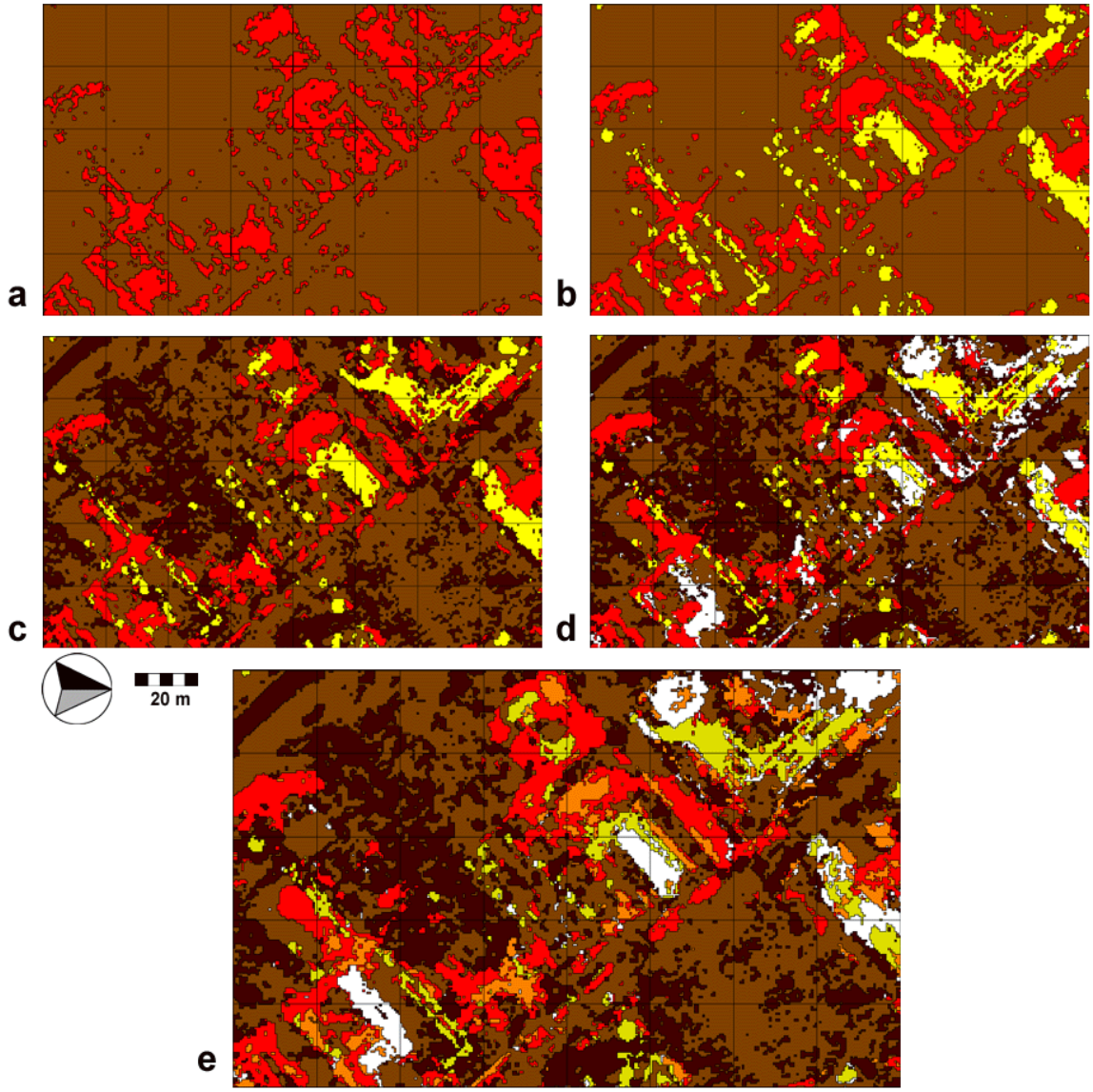


Figure 5.11. *K-means cluster analyses of the Army City data, with cluster sizes of a-e) k=2-6 classes, respectively.*

5.5.8. Evaluation

For the most part, discrete integrating methods are easy to understand and implement. Most are based on simple mathematical or GIS operations that avoid complex statistical manipulations that are difficult to follow; yet they offer very powerful results. Discrete integrating methods allow application of readily available Boolean operations to any number of geophysical data sets, thus allowing integrations of high dimensionality. A shortcoming rests in the binary maps upon which these methods are based that rely on arbitrary thresholds to define significant anomalies, forcing more subtle ones must be ignored. Integrations based on polychotomous classified data allow a larger number of thresholds, however, permitting anomalies of greater subtlety to be defined. Yet, the logical conclusion of this approach—classified maps with numerous thresholds—approaches the quality of continuous data (as does the nature of the integrated results),

suggesting that turning directly to continuous data may be a superior course of action. Perhaps most powerful are cluster analyses methods because, mathematically, they are based on multivariate mathematics that consider relationships within a multidimensional measurement space, unlike other discrete methods which are inherently one-dimensional.

5.6. CONTINUOUS INTEGRATING METHODS: SIMPLE MATHEMATICAL APPROACHES

Kenneth L. Kvamme, University of Arkansas

Continuous measurements are the raw data of geophysics, generally expressed in the domain of real numbers (or a very large range of integers). Continuous data are naturally richer than categorized information, potentially enabling superior data integrations. A small amount of prior work has been conducted in this area using a few basic techniques, but no serious investigation has yet been performed with archaeological remote sensing data. Neubauer and Eder-Hinterleitner (1997) examined sums, products, ratios, and differences between normalized resistivity and magnetometry results. Piro et al. (2000) investigated normalized sums and products of three geophysical data sets. These and more advanced methods are examined here.

Relatively simple mathematical transformations offer a rich suite of methods for integrating any number of geophysical data sets. While discrete methods based on binary data offer the advantage of clear-cut maps of anomaly presence or absence, continuous data integrations allow robust and subtle anomalies to be simultaneously expressed, producing composite imagery with higher information content.

5.6.1. Data Sum

The sum of the standardized data illustrates anomalies from all sources simultaneously, including those of large and small magnitude (Figure 5.12a). Much like the previous binary results, strong indications of foundations, floors, cellars, pipelines, and street gutters are seen throughout the town's core, but a plethora of subtle anomalies may also be discerned that helped to "fill out" entire buildings, add partitions to structures, and complete the outlines of streets, producing a richer outcome. Given the simplicity of this operation the result is particularly pleasing because it compares closely with outcomes derived by more complex means below. Variations of this method are possible by assigning more or less weight to individual sensors. For example, one might want to emphasize the resistivity contribution in defining floors and walls (Figure 5.2e). In this case, adding "2 x resistivity" to the sum would achieve this end. Different weights might be assigned to each of the other layers, offering endless permutations. Objectively, there is little reason aside from personal preferences in visualization that would justify this course of action, so it is not pursued here.

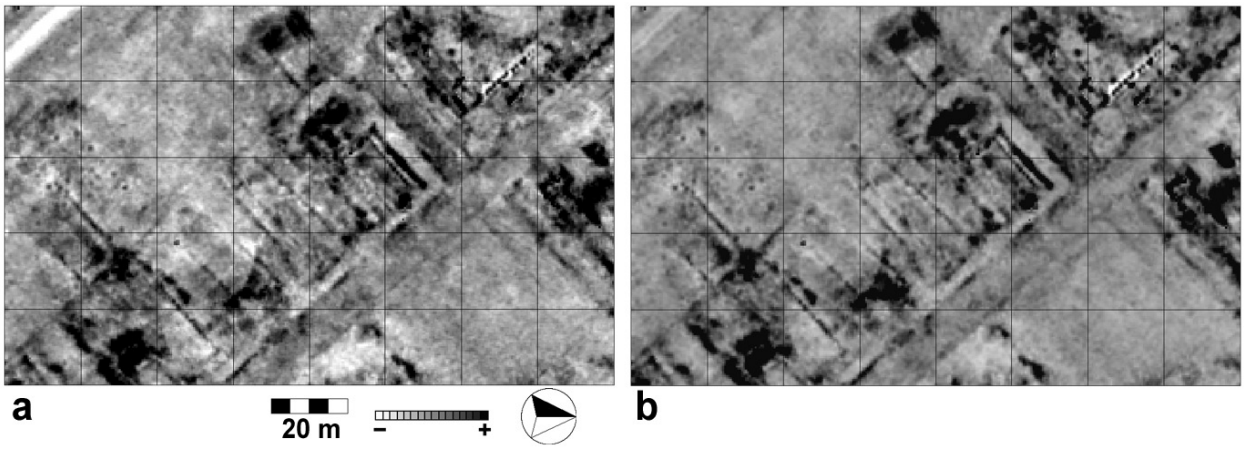


Figure 5.12. *Continuous data integrations. a) Sum. b) Product.*

5.6.2. Data Product

The data product was also investigated under the theory that cross-multiplication should make anomalies of extreme value more pronounced. A constant of 10 was first added to each standardized data set (resulting in μ=10 and σ=1) to insure positive distributions (and avoid problems caused by products of negative numbers), and the product of the six data sets was obtained (Figure 5.12b). Although the result appears broadly similar to the data sum, close inspection reveals that major anomalies tend to be somewhat better contrasted and slightly exaggerated with subtler anomalies more subdued or absent.

5.6.3. Data Ratios

The ratio between two data sets can be informative because it allows one to better contrast or highlight similarities and differences between them (Neubauer and Eder-Hinterleitner 1997). Yet, this approach is capable of integrating information from only two data sets, limiting its usefulness. For completeness, a number of data ratios are investigated here. As before, a constant of 10 was first added to each standardized data set (resulting in μ=10 and σ=1) to insure positive distributions (and avoid problems caused by negative numbers). The ratio of magnetic susceptibility to magnetic gradiometry is shown in Figure 5.13a, which indicates anomalies that occur only in the former as black, anomalies occurring only in the latter in white, and common anomalies and neutral areas as medium gray. Broad areas of shallow susceptibility in streets and around buildings are indicated (in black), while deeply buried pipes and metal artifacts detected by magnetic gradiometry are also highlighted (white). A similar result occurs in the ratio of conductivity to resistivity (Figure 5.13b), where the metallic pipes shown only by the former (in black) are contrasted against floors, walls, and footings revealed only by the later (white). Given that both of these data types are theoretically related (see Section 4.1), these images emphasize the very different contributions they make to the knowledge-base of Army City.

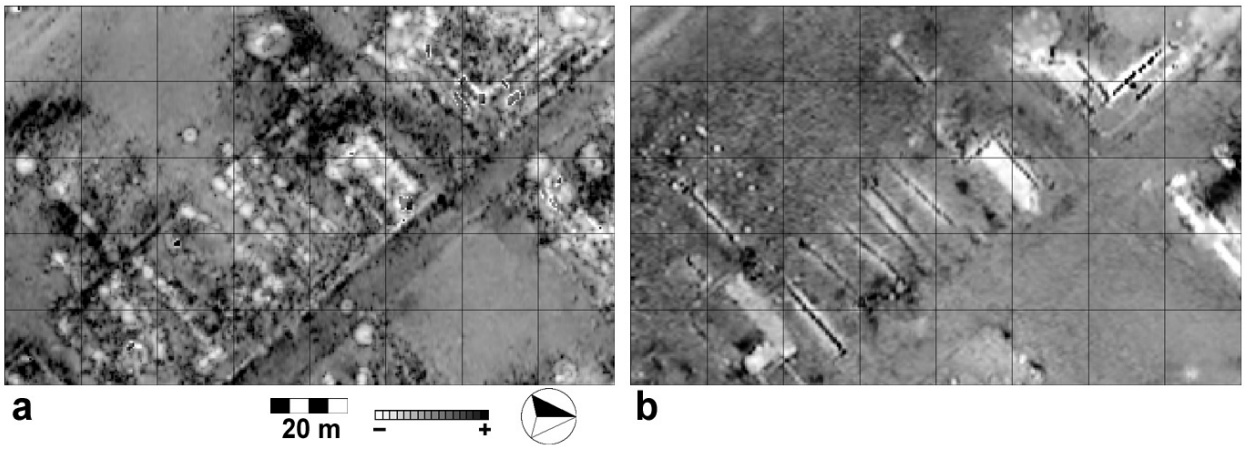


Figure 5.13. *Data ratios. a) Magnetic susceptibility / magnetic gradiometry. b) Conductivity / resistivity.*

5.6.4. Data Maximum

The *maximum* function is a tool available in most GIS (Burrough and McDonnell 1998). A per-pixel maximum of data inputs should theoretically map the largest positive anomalies in a multidimensional collection of data. At any locus (pixel), assuming commensurate measurement scales, the largest value in all inputs is taken. In other words, $x_{out} = \mathrm{MAX}(\, x_1, x_2, \ldots, x_k \,)$, where x_{out} is the transformed result, x_1 - x_k are the k input data sets, and "MAX" is the maximum function. Similar measurement scales, achievable through standardization, are necessary if each variable is to receive equal weight. A similar operation can be taken for negative anomalies (using a "minimum" function) or the maximum of absolute values might be considered. At Army City, negative anomalies were eliminated by pre-processing of both magnetic data sets (see Section 5.2.2), strong negative anomalies are nearly absent in the resistivity data (due to drought conditions at the time of acquisition; Figure 5.2e), and the same may be said of positive anomalies in EM conductivity owing to their theoretical relationship with resistivity (Figure 5.2a). Only the thermal infrared data expressed significant positive and negative anomalies (Figure 5.2f), so absolute values of this data set were computed and the conductivity data were again inverted.[2] The *maximum* function was then applied (Figure 5.14a) enabling definition of largest anomalous values across all six dimensions—apparently principal foundations, pipes, and street boundaries.

While theoretically sound, it was found that this procedure introduces unanticipated artifacts that detract from its usefulness. In relatively "quiet" areas with few prominent anomalies, small measurements that represent only random noise are apparently maximized, yielding a poor effect (these small "anomalies" probably derive from minor natural variations in soils and vegetation cover).

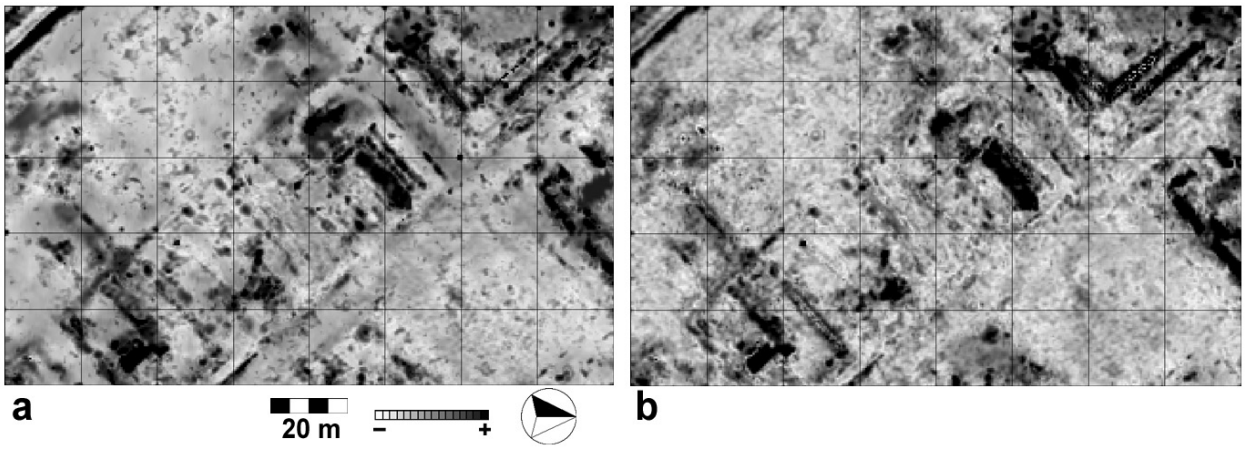


Figure 5.14. *a) Largest anomalies in the six geophysical data sets obtained through application of the* maximum *function. b) Variance of the geophysical data on a per-pixel basis, indicating anomalies that tend to be represented in relatively few data sets.*

5.6.5. Data Variance

The data integrations achieved by summing or computing the product largely reveal commonalities between the various geophysical inputs (Figure 5.12). Similarly, the binary sum and thresholds on binary sums reveal the many anomalies in common between the several geophysical methods (Figure 5.9c-e). These results beg the question of which anomalies tend to be uniquely indicated by a single geophysical method or by relatively few methods. Computing the variance, $\Sigma_i(x_i-\mu)^2/n$ (where μ is the mean of the n=6 sensor inputs in a pixel), across the six dimensions and mapping the result on a per-pixel basis allows visualization of these anomalies (Figure 5.14b). High variance areas (in black) point to regions of highly variable geophysical results across the data sets, suggesting loci where results tend to be very different between one or more data sets. A high variance might be derived where one geophysical technique indicates an anomaly, but the others do not, or where one method indicates a positive anomaly and another a negative one, for example. On the other hand, locations without anomalies in any data set, or those uniformly indicated as anomalous, will show little variance (mapped in light gray to white).

5.6.6 Evaluation

Simple mathematical approaches to data integration are extremely easy to implement and understand. They also yield effective data fusions. Sums and products appear to be particularly effective, which probably explains why they have appeared elsewhere in the limited literature that integrates sensor data (e.g., Neubauer and Eder-Hinterleitner 1997; Piro et al. 2000). Surprisingly, it will be shown later (Section 5.10) that the simple unweighted sum closely parallels results obtained by advanced multivariate techniques. Other methods explored here are shown to be less useful: the ratio is limited to only two layers, the maximum function highlights noise in uniform regions devoid of anomalies, and the variance function does not illustrate common anomalies generally revealed by different sensors.

5.7. CONTINUOUS INTEGRATING METHODS: MULTIVARIATE TECHNIQUES

Kenneth L. Kvamme, University of Arkansas

The methods of multivariate statistics offer a powerful means for data integration and they give insights about structure and interrelationships between geophysical dimensions. A host of such methods is available, and may broadly be divided into two classes: those that reduce dimensionality and classification methods. The former includes principal components and factor analyses that produce linear combinations of the original data. The significant components that result, fewer in number than the original variables, represent data fusions based on statistical correlations. Classification methods, on the other hand, attempt to define "groups" or "classes" in bodies of data through two different methods. One such approach is *cluster analysis* where observations are grouped into "natural" classes based on their proximities in a measurement space, producing a discrete result described earlier (Figure 5.11). The other includes statistically optimized *discriminant functions* derived from data patterns in samples from known classes. These continuous functions are applied to data of unknown membership to generate linear composites or probability surfaces for the defined classes. As previously noted, the former is referred to as "unsupervised classification" because the algorithm finds naturally occurring classes in multivariate data. The latter is known as "supervised classification," because functions are developed from statistical characteristics of samples provided by a supervisor (Schowengerdt 1997). It is emphasized that classification models of any kind may be regarded as having a predictive capacity when mapped classes indicate not only robust and well-recognized anomalies, but vague or nearly invisible ones as members of the same class. In these cases, such methods can be regarded as predicting from the known (anomalies of certain or nearly certain identity) to the unknown (anomalies of uncertain membership). In so doing, prospecting possibilities are augmented and a form of subsurface predictive model is derived.

Before proceeding, it is prudent to first examine a fundamental characteristic of the Army City geophysical data. It has been indicated that many anomalies do not recur in different data sets suggesting that independence or a lack of correlation may exist (compare elements of Figure 5.2 and 5.14b). The assumption of correlation lies at the very heart of such methods as principal components analysis, making this well-known method of data fusion possible. On the other hand, independence between variables is an important assumption behind many multivariate statistical models, such as multiple regression (Davis 2002). A Pearsonian correlation matrix was computed between the six geophysical data sets (Table 5.3). Noteworthy are generally low levels of correlation, a circumstance that is surprising because, theoretically, one might suspect resistivity and conductivity to be highly correlated, as well as (absolute) magnetic gradiometry and susceptibility. Yet, the highest absolute correlations in the entire data set approach only $|r|=.3$ (about $100r^2=9$ percent of variance in common), with the thermal data being almost completely independent. These data shed light on the performance of principal components and related methods, and aid in the interpretation of various results throughout this study.

Table 5.3. *Pearsonian correlation matrix between the six normalized geophysical variables. The largest off-diagonal absolute correlation in each column is indicated in bold typeface.*

	Cond	**GPR**	**Mag**	**MS**	**Res**	**Therm**
Cond	1.000	-.160	.073	-.286	-.137	**-.081**
GPR	-.160	1.000	.231	**.304**	.218	-.053
Mag	.073	.231	1.000	.299	**.277**	-.019
MS	**-.286**	**.304**	**.299**	1.000	.073	.022
Res	-.137	.218	.277	.073	1.000	-.035
Therm	-.081	-.053	-.019	.022	-.035	1.000

KEY: **Cond**=EM conductivity; **GPR**=ground-penetrating radar (maximum positive amplitude); **Mag**=absolute magnetic gradiometry; **MS**=absolute magnetic susceptibility; **Res**=electrical resistivity; **Therm**=thermal infrared

5.7.1. Principal Components Analysis

Principal components analysis (PCA) linearly combines multivariate data based on the correlation structure between *k* variables (Table 5.3). The resulting *k* components are uncorrelated and ordered in such a way that the first represents more of the total variance in the data than the second, which in turn carries more than the third, and so on (Davis 2002:509-525). When standardized, each variable contributes a variance of unity, and the first few components typically characterize the bulk of the information content in most data sets. A PCA of the Army City data reveals that only the first two components account for more variance than any single variable alone (eigenvalues larger than unity), with the first accounting for nearly 30 percent of the total variance, and the second about 19 percent (Table 5.4). These low values result from the lack of correlation between the variables (Table 5.3). Yet, that the first two components account for nearly half the total variance offers encouragement that useful fusions of data commonalities is achieved. The *loadings* (correlations of variables with a component) on the first component indicate it to be primarily a linear combination of *GPR*, *magnetic gradiometry, susceptibility*, and *resistivity*, with moderate contributions from each (Table 5.4; Figure 5.15a). The second component primarily reflects *conductivity*, with a lesser contribution from *thermal infrared* (Table 5.4; Figure 5.15a). The remaining components more or less correspond with individual variables, with the third represented largely by *thermal infrared* (Table 5.4; Figure 5.15c), the fourth by *resistivity* (Table 5.4; Figure 5.15d), and the fifth by *GPR* (Table 5.4; Figure 5.15e). The sixth or last component, accounting for only eight percent of the total variance, forms a composite that is not dominated by a single variable, but one that shows roughly equal loadings among four variables (*conductivity*, *gradiometry*, *susceptibility*, *resistivity*; Table 5.4; Figure 5.15f). With the lower components usually regarded as containing left over and idiosyncratic noise, this component appears to well represent larger metallic artifacts, including the many pipes that cross the site that are unrelated to the other dimensions.

Table 5.4. *Principal component analysis of the standardized geophysical data. High loadings in columns are indicated in bold typeface.*

Component	PCA Eigenvalues		
	Variance	**% Variance**	**Cumulative %**
1	1.791	29.86	29.86
2	1.153	19.23	49.07
3	.964	16.06	65.13
4	.899	14.99	80.12
5	.714	11.91	92.02
6	.479	7.98	100.00

	PCA Loadings (correlations with components 1-6)					
	1	**2**	**3**	**4**	**5**	**6**
Conductivity	-.423	**.701**	.313	.284	.219	.321
GPR	**.681**	.042	-.169	.073	**.704**	-.072
Magnetic gradiometry	**.611**	.467	.388	.217	-.259	-.379
Magnetic susceptibility	**.694**	-.258	-.060	.477	-.249	.397
Resistivity	**.540**	.280	.130	**-.732**	-.095	.262
Thermal infrared	-.021	**-.545**	**.816**	-.045	.182	.030

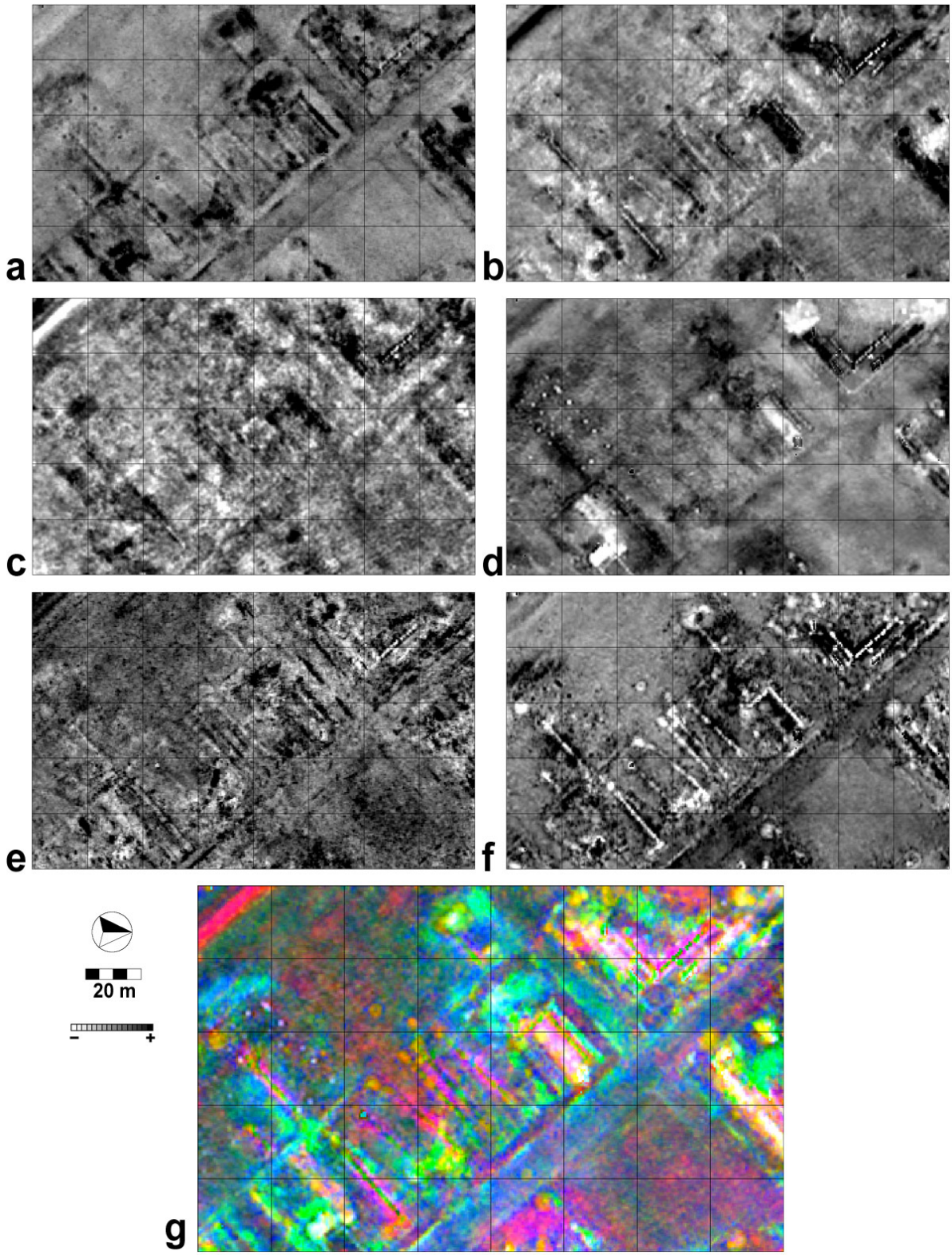


Figure 5.15. *Mapped results of principal component analyses. a-f) Components 1-6, respectively. g) RGB color composite of components 1 (green), 2 (red), and 6 (blue).*

Components 1, 2, and 6 each represent linear composites of high dimensionality, or excellent fusions of the geophysical data. All are uncorrelated, however, and therefore each represents an integration of a *different* fundamental characteristic of Army City's subsurface. Earlier compositing methods may be made use of here to integrate these individual results into a single complete whole. An RGB color composite of these components is offered in Figure 5.15g, with Component 1 assigned to green, Component 2 to red, and Component 6 to blue. This result, accounting for 57 percent of the total variance in the data, well reveals roads, pipes, floors, wall, footing, gutters, and more, forming one of the best data fusions thus far achieved. The next section attempts to improve on this result by turning to factor analysis methods.

5.7.2. Factor Analysis

While the foregoing principal components are informative, the variance they represent might be better distributed between the components through a *varimax* rotation, sometimes referred to as factor analysis (Davis 2002:533). This method undertakes an orthogonal rotation of the significant components (with Eigenvalues greater than unity—Components 1-2) within the six-dimensional measurement space to more equally distribute the variance they represent. In so doing, the loadings with the original variables are altered as variables with high loadings are more equitably distributed across both factors. The outcome is that the variables represented in each component can change resulting in improved interpretability. This was undertaken and, while the variance accounted for by the second component or factor after rotation is only slightly increased from 19.2 to 20.5 percent (with a corresponding decrease in Component 1), the loading structure is changed yielding somewhat clearer results (Table 5.5). The contribution of *conductivity* is nearly removed in the first rotated factor and increased in the second, but inverted (i.e., a negative correlation). The influence of *magnetic gradiometry* is greatly reduced in the second component, although *magnetic susceptibility* makes a larger contribution (compare Tables 4.4 and 4.5). The first factor represents a linear composite of *GPR*, *magnetic gradiometry*, *magnetic susceptibility*, and *electrical resistivity*. The second factor combines *conductivity* and *thermal infrared* with some influence from susceptibility (Table 5.5).

Table 5.5. *Factor analysis of the geophysical data, with* varimax *rotation. Principal loadings in columns are indicated in bold typeface.*

Factor	Factors after Rotation		
	Variance	% Variance	Cumulative %
1	1.712	28.54	28.54
2	1.232	20.53	49.07

	Factor Loadings (correlations with factors) after Rotation		Factor Score Coefficients	
	1	**2**	**1**	**2**
Conductivity	-.150	**-.805**	-.007	-.652
GPR	**.652**	.210	.369	.100
Magnetic grad.	**.737**	-.222	.462	-.259
Susceptibility	**.559**	.486	.284	.346
Resistivity	**.604**	-.072	.367	-.121
Thermal	-.212	**.503**	-.177	.439

These uncorrelated factors are illustrated in Figure 5.16a, b and represent independent underlying dimensions of the Army City subsurface, accounting for 49 percent of the total variation. The first points to the foundations, floors, walls and gutters revealed by magnetics, resistivity, and GPR. The second emphasizes such negative anomalies as pipelines, probable cellars, and a modern drainage ditch (near the southwest corner of the study area) that tend to be seen in the conductivity data. These results appear broadly similar to the first two principal components (compare Figure 5.15a,b), although many subtle differences occur throughout both maps and the negative correlation of *conductivity* causes the gray scale to be inverted relative to PCA Component 2. Factors 1 and 2 are combined in a blue-green color composite in Figure 5.16c. It is of interest that nearly every anomaly seen in the three-color PCA composite of Figure 5.15g appears in the two-color factor composite of Figure 5.16c, probably because the third component of the former accounts for only an insignificant amount of the total variance (less than eight percent).

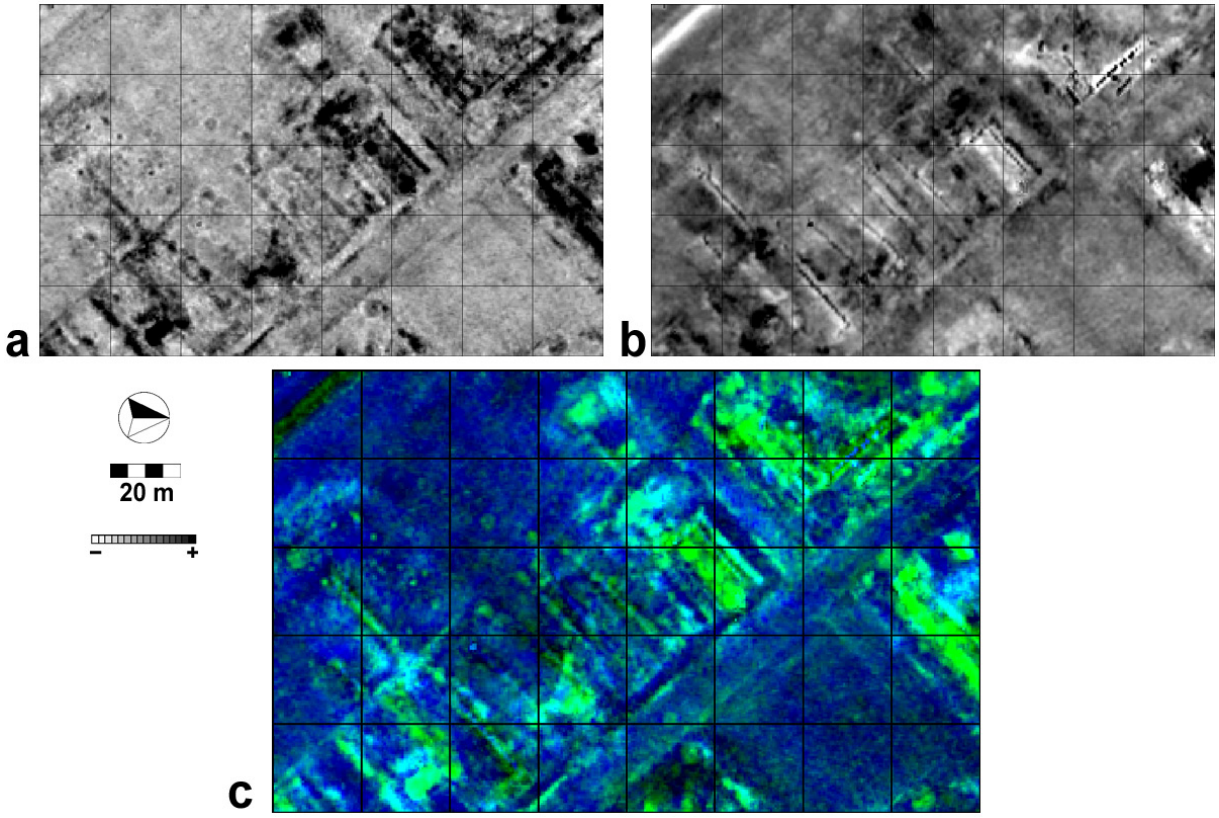


Figure 5.16. *Factor analysis of the Army City data. a) Factor 1. b) Factor 2. c) Color composite of both factors with Factor 1 assigned blue, Factor 2 assigned green.*

5.7.3. Supervised Statistical Classifiers

A large number of algorithms exist in satellite remote sensing for developing supervised classifications. The classical methods explored here are pixel-based where the unit of analysis is the pixel and its multivariate characteristics across several bands or data layers. Newly developed methods, investigated below, also consider context and data patterns in local neighborhoods (Jensen 2005). Many pixel-based methods utilize a multivariate normal model (e.g., maximum likelihood, discriminant functions), while others apply simple geometric operations to the multidimensional measurement space (e.g., minimum distance and parallelepiped classifiers; Schowengerdt 1997). Illustration of these many methods would be fruitless because of broadly parallel results, but two case studies are examined to illustrate their potential for data integration. A logistic regression algorithm is investigated because it represents a particularly robust nonparametric classifier (Hosmer and Lemeshow 2000; see Section 5.11 for a comparison of this method with linear discriminant functions). As a start-point, a simple binary model is pursued for robust anomaly presence or absence. It is then succeeded by a multi-class solution that develops models for several specific *types* of anomalies based on known excavation results (described in Section 5.8). It is emphasized that these models, which yield continuous results interpretable as probability surfaces for anomaly presence, also possess a *predictive capacity*, because by consolidating the multidimensional data they may indicate previously unrealized anomalous locations that

do not exist in the training data and which may be unseen in the individual geophysical data layers from which they are derived.

5.7.3.1. Binary logistic regression

The selection of "training sites"—locations of known class membership provided to the algorithm—is vital to a good result because the resulting classification function is optimized to patterns in these samples. With a goal of integrating the six-dimensional geophysical data into a one-dimensional composite that characterizes anomaly distributions, training site selection was investigated using a number of approaches. One that works well is a model of robust anomalies, as represented by the earlier binary result (Figure 5.9d) that shows anomalies indicated by at least two techniques (re-illustrated in Figure 5.17a). Such a model could tell us more about strong anomaly occurrences by generalizing patterns in the six-dimensional data. Fifty percent of the cases were drawn at random for development of the model (32,000 data points). This is a common tactic in supervised classification because model parameters fit to a particular sample generally may not perform as well when classifying other independent samples (Schowengerdt 1997). The 32,000 data points *not* used for model development are then used below for a truer assessment of model performance in classification. No interaction terms were selected and all variables were forced into the model, resulting in *pseudo*-R^2=.63. The following function was derived that maximizes differences between the two classes (i.e., the presence or absence of a robust anomaly):

$$L = -2.846 - .171 \times Cond + 1.101 \times GPR + .974 \times Mag + .859 \times MS + .757 \times Res - .436 \times Therm$$

(see Table 5.3 for explanation of abbreviations). All parameters in this model are statistically significant contributors (at $\alpha < .001$) and they offer interpretive potential. With the data standardized the absolute sizes of the coefficients indicate that GPR and magnetic gradiometry contribute most to discriminating robustly anomalous locations from the background, conductivity and thermal contribute least, and the signs inform us that negative values of conductivity and thermal but positive values of the other variables are most related to anomaly presence. The logistic transformation:

$$p = (1 + exp(-L))^{-1}$$

conveniently rescales this axis to a 0-1 range, which may be interpreted as a probability surface for robust anomaly presence (Figure 5.17b). This map well represents the anomalies of the training sites (Figure 5.17a), but also extends them beyond their original limits. In doing so, other locations of robust anomalies may be indicated, including some previously illustrated and some not yet realized. The result clearly forms an effective fusion of all the geophysical data.

It is emphasized that the training data used for this model are themselves a model, based on all pixels indicated as anomalous by two or more geophysical methods (Figure 5.17a). Assessing the accuracy of the regression model in mimicking this other model does not therefore carry the same weight as a comparison against true archaeological circumstances (as is done in the following section). Nevertheless, it is worthwhile to examine its performance against the reserved test sample, because it indicates anomalous areas beyond the discrete boundaries of the training data and gives insights about other

anomalous locations. This is accomplished by selecting a threshold on the *p*-value axis in order to produce a binary model for anomaly presence or absence. With approximately 16 percent of the training data indicated as robustly anomalous (Figure 5.17a), a *p*-value threshold of *p*=.359 was determined in the probability surface of Figure 5.17b, above which about 16 percent of the model's pixels could be discretely mapped as anomalous (Figure 5.17c). Cross-tabulating this result against the test sample yields a "confusion matrix" that allows quantitative assessment of model performance (Table 5.6a). These data indicate that approximately 72 percent of the anomalous pixels in the test sample are correctly specified, as are about 95 percent of background pixels lacking anomalies, with an overall accuracy of 91 percent.

Table 5.6. *Confusion matrices showing performance of the binary logistic regression model against robust anomalous locations in the test sample. a) Model at a p=.359 threshold. b) Model at the p=.153 threshold. Column percents given in table.*

a		**"TRUE" CATEGORY: Anomalies indicated by 2 or more methods (Figure 5.17a)**		
		0	**1**	**Total**
PREDICTED CATEGORY by regression model p>.359 (Figure 5.17c)	**0**	25,535 (94.85%)	1425 (28.06%)	26,960 (84.25%)
	1	1387 (5.15%)	3653 (71.94%)	5040 (15.75%)
	Total	26,922 (84.13%)	5078 (15.87%)	32,000
Overall Kappa 0.67; Overall accuracy: 91.21%; Average accuracy: 83.40%				

b		**"TRUE" CATEGORY: Anomalies indicated by 2 or more methods** (Figure 5.17a)		
		0	**1**	**Total**
PREDICTED CATEGORY by regression model p>.153 (Figure 5.17d)	**0**	23,488 (87.24%)	762 (15.00%)	24,250 (75.78%)
	1	3434 (12.76%)	4316 (85.00%)	7750 (24.22%)
	Total	26,922 (84.13%)	5078 (15.87%)	32,000
Overall Kappa 0.60; Overall accuracy: 86.89%; Average accuracy: 86.12%				

While this performance represents a very good overall fit, its accuracy in classifying the more important anomaly class is not outstanding, with an error of omission of 28 percent (Table 5.6a). It should be noted that the figures in this table are somewhat arbitrary, however. A probability surface is a continuum; by selecting a lower *p*-value threshold a larger region can be mapped to the anomaly class causing more truly anomalous regions to be captured. Although this will reduce errors of omission it does so at the price of more errors of commission—a useful tradeoff when one class is of greater importance. This was undertaken by selecting a percent correct target of 85 percent of the anomalies in the test sample, which was found to be associated with a probability threshold of *p*=.153. The resultant binary map is also illustrated in Figure 5.17c. Cross-

tabulating this model's predictions against the test sample yields performance data given in Table 5.6b. While 85 percent of true robust anomalies are correctly specified, the cost is a model mapping that covers nearly a quarter of the study area (compared to less than 16 percent in the previous classification of Table 5.6a). Overall accuracy is reduced to about 87 percent.

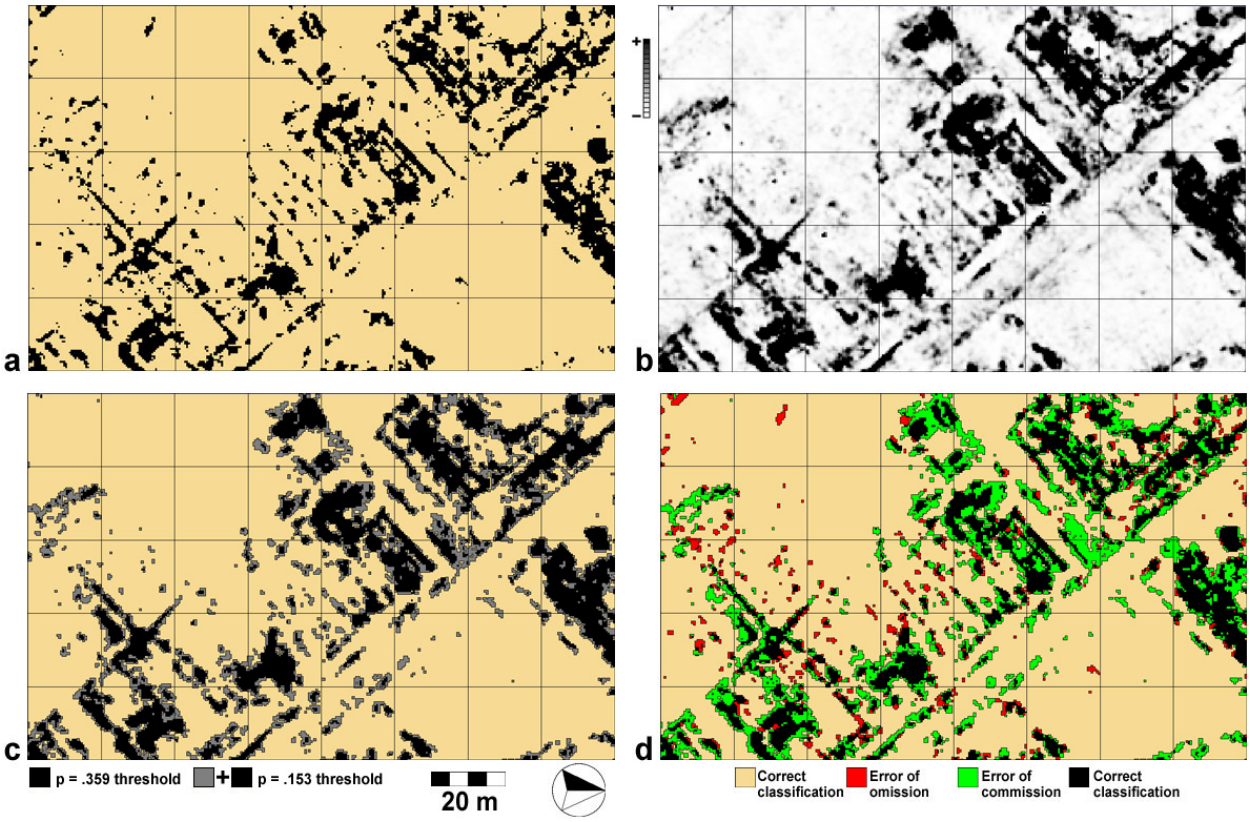


Figure 5.17. *Performance of the binary logistic regression model. a) Robust anomalies indicated by two or more geophysical methods from which training samples were drawn. b) Logistic regression probability surface for robust anomaly presence. c) Classification of the regression model based on* p=*.359 and* p=*.153 thresholds. d) Distributions of model errors for the* p=*.153 classification.*

The distributions of both types of model errors for the *p*=.153 threshold are mapped in Figure 5.17d, which improves understanding of model performance. Errors of commission occur through model over-specification, where locations truly belonging to Class 0 are assigned to Class 1. They are indicated in green and include about 13 percent of the actual Class 0 pixels (Table 5.6b). Comparison against the probability surface (Figure 5.17b) and previous integrations strongly suggests they represent true anomalies that extend beyond the limits of the discrete mapping of the training class because they generally occur adjacent to these anomalies. In other words, they probably do not indicate real errors. Errors of omission, on the other hand, are more serious because they point to potential anomalies missed by the model. They occur when pixels truly belonging to Class 1 are erroneously assigned to Class 0, which now includes only 15 percent of the Class 1 pixels (about 2.4 percent of the total pixels; Table 5.6b). They are shown in red in

Figure 5.17d and tend to indicate smaller and isolated point anomalies as well as ones associated with pipes. Comparison against the probability surface (Figure 5.17b) suggests that an even lower threshold could capture many of these anomalies, but this would be accomplished only by trading an even higher rate of errors of commission.

5.7.3.2. Multinomial logistic regression

In this section a multiple-class logistic regression model for specific *types* of anomalies is developed based on known excavation results (described in Section 5.8). Consequently, it is very different from the preceding example because it attempts to model true archaeological circumstances revealed by excavation. As such, it represents a "predictive archaeological model" for the subsurface, and therefore offers yet another applications area developed by this project. This model will statistically combine six-dimensional geophysical information by correlating the data with known archaeological finds. It will then probabilistically project similar circumstances throughout the Army City study area. In so doing, it will indicate loci beneath the surface with conditions identical or very similar to the known archaeological classes, thereby yielding a new form of archaeological prospecting that should augment current prospecting capabilities. In other words, the predictive model, by combining data from *all* sensors, may point to subsurface archaeological features with greater accuracy than can any single sensor alone. The outcomes, of course, represent high-dimensional data fusions.

In the archaeological testing and validation phase, described in Section 5.12.3.1, over 100 small excavations were made throughout Army City in an effort to evaluate anomalies indicated by the geophysical surveys and assess the accuracy of sub-surface predictions. Excavation units were placed by random sampling, stratified by general classes of anomaly types that included a "background" class devoid of anomalous indications. Twenty-two of the excavation units identified buried concrete or masonry representing floors, walls, and footings associated with former structures. Another nine units revealed pipes and 16 cross former street or alley gutters. These classes are the focus of this model, but two additional categories are necessary. One is a single class representing all other anomalous locations representing archaeological finds in 26 excavation units (i.e., archaeological features that represent neither concrete/masonry nor pipes nor gutters). Additionally, a reference class is statistically required against which characteristics of these four archaeological classes can be contrasted. Eleven excavation units from the background class represent loci without archaeological evidence or anomalous geophysical indications.

To summarize, the archaeological and background classes include: (1) concrete/masonry rubble, (2) pipes or large metal objects, (3) gutters, (4) other archaeological features grouped collectively, and (5) the background devoid of archaeological evidence. Training samples for the archaeological classes were obtained from all anomalous areas within the excavation units and adjacent areas of the *same* anomalies outside the units; the background class utilized the areas within a two-meter radius of the background excavations. Training locations are indicated in Figure 5.18a.

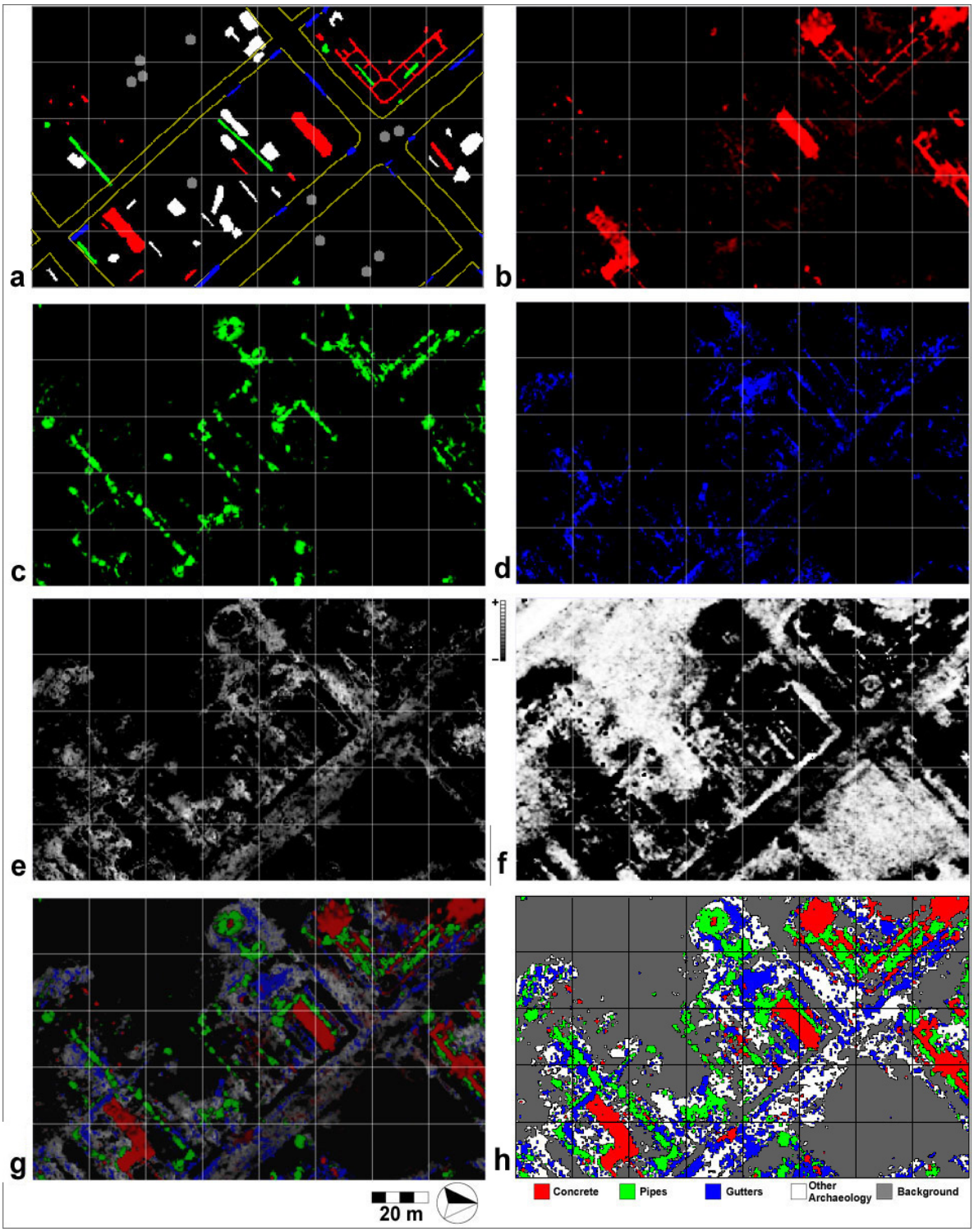


Figure 5.18. *Modeling specific archaeological classes with multinomial logistic regression for data fusion. a) Training sites with roads and alleys indicated. b) Probability surface for concrete/masonry. c) Probability surface for pipes. d) Probability surface for street gutters. e) Probability surface for other archaeological features. f) Probability surface for anomaly absence. g) Color composite of combined probability surfaces. h) Maximum likelihood classification of a-f.*

With multiple classes, the goal is to develop a probability surface for the presence of *each* one, requiring a multinomial model. Although fuzzy logic (Jensen 2005:189) could be employed where a pixel might be classified as 35 percent pipeline and 65 percent concrete, given the imprecision of geophysical and archaeological data no attempt is made to do so. All classes are treated as mutually exclusive and the regression results as "hard classifications." The training data of Figure 5.18a are employed with the five indicated classes as the dependent variable and the six geophysical data sets as independent variables. For simplicity, equal class a priori probabilities are assumed, allowing results that are solely a function of geophysical data patterns rather than the presumed or estimated prevalence of classes in the study region. Finally, split sampling was employed where the models are developed from 75 percent of the training data, and tested independently on the remaining "test" sample of 25 percent, for a truer indication of model classification performance.

With the background ("no archaeology") class serving as the referent, four regression functions are obtained (Table 5.7), where for each class, k :

$$L_k = \alpha_k + \Sigma_{i=1,6}\beta_{ki}X_i$$

is applied to each pixel. With standardized data, these functions themselves offer considerable interpretive potential through comparisons between the coefficients. *magnetic gradiometry* and *susceptibility* most influence the "pipes" model; *magnetic susceptibility*, *resistivity*, and *GPR* best define street gutters; *resistivity* and *magnetic susceptibility* carry most weight in the "concrete-masonry" model and in the "Other archaeology" model. These findings agree well with data patterns in the initial geophysical imagery (Figure 5.2) and make geophysical sense. Concrete and masonry are high resistance features, iron pipes possess a high level of induced and native magnetism, in-washed topsoil filling street gutters tends to be more magnetically susceptible, and the rubble within them exhibits high resistivity and possesses dielectric differences for GPR reflections.

Table 5.7. *Parameter estimates for the four logistic regression functions based on 75% of the training data.*

	MODEL CLASS (k)			
PARAMETERS	**Pipes**	**Gutters**	**Concrete**	**Other**
Intercept (α)	-.588	-.142	-1.442	.451
Conductivity (β_1)	.177*	.557	.796	.186*
GPR (β_2)	.587	1.375	1.010	.891
Magnetometry (β_3)	2.038	-.047*	.421	.636
Mag.Susceptibility (β_4)	2.701	3.373	2.965	2.997
Resistivity (β_5)	.604	1.936	3.480	1.703
Thermal (β_6)	.196*	.194	.466	.294
Sample sizes (N=3290)	302	351	1072	1173

*Parameters not significant at α=.10

A probability surface for any class is generated by:

$$p_k = \exp(L_k) / [\ 1 + \Sigma_i \exp(L_i)\ ]$$

where p_k is the probability of membership in class k, L_k is the linear regression function for that class (Table 5.7), "exp" is the exponential function, and the summation is over all classes (Hosmer and Lemeshow 2000). When mapped on a per-pixel basis through GIS map algebra methods, a probability surface for each class is obtained. A probability surface for the fifth or background class is obtained by subtracting the sum of the four class probabilities from unity.

The five probability models are portrayed in Figure 5.18b-f, and simultaneously in Figure 5.18g though use of color compositing (see Section 5.4.2). These probability surfaces represent prospecting models for specific archaeological classes lying beneath the surface—pipes, street gutters, concrete-masonry, and other archaeological features. Visual comparison against the training sites suggests very good correspondence. It is emphasized that the model for the background class, if inverted (i.e., subtracted from unity), itself represents a good fusion model for all archaeological classes considered together, comparable to the binary model of the preceding section. The success of these models arises because they are based on multidimensional geophysical data that combine information from all inputs. Each model contains a predictive element because it maps localities beyond those indicated in the training data (Figure 5.18a). In other words, locations are identified that possess characteristics similar to the known archaeological classes that might not be seen as strongly anomalous in the original input data (Figure 5.2). In this way, the GIS-derived models become predictive because they extend the prospecting capabilities of geophysics by indicating potentially new and unknown anomalous locations associated with specific classes of archaeological features.

While good performance is suggested by the mappings, these models do not perform perfectly because each probability surface points to locations *not* of the appropriate class. These errors can be assessed quantitatively by assigning each training pixel to its highest probability class of membership to form a classified map for the five classes (Figure 5.18h). This result may then be cross-tabulated with the training data for a classical per-pixel performance evaluation (Table 5.8). Although the overall accuracy is only about 67 percent, and class accuracies vary between 41-92 percent, it is clear that for each actual class the models assign the majority of pixels to the correct class. Most of the errors appear to be associated with the "other" class, but that is to be expected. That class, after all, holds a wide variety of anomalies representing many types of archaeological features, including subtle walls, floors, street components, burned areas, and other disturbances similar to the principal classes. It is therefore unsurprising that the variation it represents appears very similar to them. Without this class, the overall accuracy raises to an impressive 82 percent. The background, pipe, and concrete-masonry models are most accurate, but the gutters model is not robust. This is probably reasonable because in the primary data (Figure 5.2) it is apparent that gutters generally exhibit only modest anomalies with mid-range measurements very similar to other anomalies, making their discrimination difficult. Errors of commission (100 minus row percent correct) are generally high, however, with the gutters class particularly over-specified, indicating that other archaeological features geophysically look like gutters. The concrete class performs

particularly well with the lowest errors of commission, but this anomaly type appears most robust and discrete in the raw data (Figure 5.2). Again, most of errors of commission derive from the "other" class.

Table 5.8. *Classification accuracies of the Army City models applied to the 25 percent test sample of known archaeological circumstances (Taken from Figure 5.18a).*

	TRUE CLASS						
PREDICTED CLASS	**Background**	**Pipes**	**Gutters**	**Concrete**	**Other**	**TOTAL**	**% correct**
Background	**105**	2	1	5	50	**163**	64.4
Pipes	0	**102**	9	11	40	**162**	63.0
Gutters	0	8	**56**	23	97	**184**	30.4
Concrete	0	1	4	**306**	55	**366**	83.6
Other	9	7	28	13	**165**	**222**	74.3
TOTAL	**114**	**120**	**98**	**358**	**407**	**1097**	
% correct	92.1	85.0	57.1	85.5	40.5		

Overall accuracy: 66.9%; Average accuracy (columns): 72.0%

It is emphasized that the confusion matrix in Table 5.8 represents model accuracy on a pixel-by-pixel basis. The pipes model, for example, might correctly specify only 25 of the 50 pixels (50 percent) composing a known pipe. Looked at on a per-feature basis, however, with half a pipe's pixels indicated by a model in a linear arrangement, it is likely an analyst would correctly identify the feature (in fact, a linear pipe or wall feature might be identified by perhaps only 20 percent of its pixels—or at least warrant investigation)! Viewed in this way, per-pixel performance evaluations grossly underestimate true model performance in practice. Nevertheless, great promise for improved prospecting is suggested by these results and additional work is clearly necessary to refine this approach. The following sections explore recently developed and advanced supervised classification algorithms for doing so that yield alternative integration approaches.

5.7.6. Evaluation

Multivariate statistical methods are extremely varied conceptually and in the results they offer. All yield powerful and effective data integrations. On top of this, principal components and factor analyses offer graphical fusion capabilities when components are composites of color. All methods offer important quantitative insights in the form of coefficients, loadings, or other statistics that lead to improved theoretical knowledge of relationships and greater interpretive capabilities. When training data are available, supervised classification methods allow the modeling of specific classes of archaeological features and therefore models of great specificity. These models may actually extend prospecting capabilities because in the fused results they may reveal the loci of important features not indicated by any single sensing method.

5.8. INTELLIGENT KNOWLEDGE-BASED SYSTEMS FOR DATA INTEGRATION

Jason A. Tullis, University of Arkansas

5.8.1. Artificial Intelligence and Geocomputational Tools

Propelled by its short history of approximately 50 years, artificially intelligent (AI) computing now offers an expanding variety of techniques for extracting information from remote sensor and ancillary spatially distributed archaeological datasets. A mid-twentieth century test for computer intelligence required that a machine's response to an interrogator's question be indistinguishable from a human being's response at least 30 percent of the time (Turing 1950) – quite a difficult challenge even today! Like many other kinds of applications in AI, *geocomputation* (a term for the emerging field of AI-assisted computational geomatics) is not interested in passing a "Turing Test," but in software that extracts information from data (or performs other data-related tasks) in a manner *better* than human beings can do it. Given this broad definition, most software algorithms have some level of *machine intelligence*, however small the amount (Openshaw and Openshaw 1997). Current geocomputational tools include a variety of AI-based techniques, such as expert systems, neural networks, machine learning decision trees, genetic algorithms, and agent-based systems (Jensen 2005; O'Sullivan and Unwin 2003).

Machine learning decision trees have gained significant momentum in applications of geocomputation (e.g., Jensen 2005; Hodgson et al. 2003; Tullis and Jensen 2003; Muchoney et al. 2000), largely due to their simplicity and speed for predicting numbers (regression) or classes (classification) from example data (analogous to training sites, as in Section 5.7.3.2). On a higher level of complexity in configuration, implementation, and logic, *neural networks* can also regress and classify unknown cases after being properly trained from example data. They are capable of learning more complex patterns, but are significantly slower than machine learning decision trees. *Genetic algorithms* are based on the concept of evolution (changes in the frequency of an allele over time) and improve using modeled processes of *crossover*, *mutation* and many *new generations* of the algorithm. The future effectiveness of genetic algorithms, neural networks, machine learning decision trees, and other geocomputational techniques will ultimately rest on higher levels of intelligent processing that control these and other tools. *Intelligent agent* technology is gaining attention because it offers hope for *autonomy*, *reactivity*, and *goal direction* (O'Sullivan and Unwin 2003). Imagine an information extraction algorithm with these characteristics that can make decisions on its own, evolve, and learn. Much work is required to bring these ideas to an operational level in geocomputation.

State-of-the-art machine learning decision trees were selected to define subterranean archaeological features in Army City by integrating its geophysical data sets. The reason for this selection parallels the expanding interest in these methods in other information extraction applications (e.g., medical diagnosis). While lacking some of the pattern recognition capabilities of neural networks (which can automatically encode index-like patterns among multiple variables), machine learning decision trees offer speed, flexibility, and *intelligible* rules (i.e., a series of "if-then" statements). They are far less complicated to deploy and parameterize, and offer competitive accuracy in many

applications where the two are compared. The other geocomputational techniques mentioned, including genetic algorithms and agent-based systems, were not investigated in this study but are on the horizon for this type of archaeological information extraction.

5.8.2. C5.0 and Machine Learning Decision Trees

Jensen (2005) provides an excellent summary of machine learning decision trees. He first explores the concept of *expert systems*, which carry a relatively long history within AI. Expert systems are powerful because they capture and use knowledge from *human* experts. They are used extensively in the world today (e.g., automotive fuel injection systems, anti-lock brakes, automated flight control systems, satellite guidance systems). Their operation is different from that of conventional computer algorithms that are embedded with human knowledge. In expert systems, human (expert) knowledge is stored in a database (e.g., as production rules) where it can be retrieved to solve a specific problem. Expert systems are heuristic (involving feedback loops, rules-of-thumb, etc.), increasing the level of machine intelligence, broad as that definition may be. In traditional expert systems the expert explicitly defines rules. In some cases, semi-automated techniques are utilized to extract rules from the expert. In the case of machine learning decision trees, the rules can be *automatically* extracted by induction (from examples to production rules).

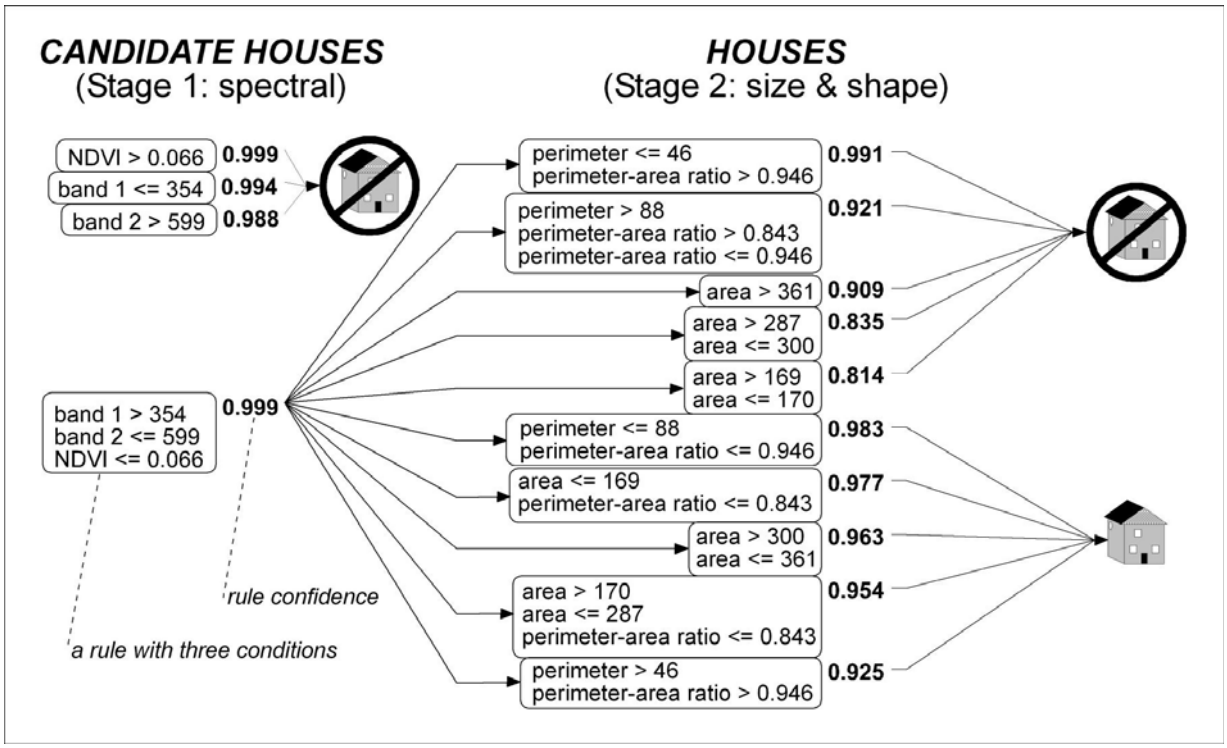


Figure 5.19. *C5.0 rules generated for the extraction of residential dwellings from pan-sharpened IKONOS imagery (adapted from Tullis and Jensen 2003).*

Tullis and Jensen (2003) utilized the C5.0 (called *See5*) machine learning decision tree algorithm, developed by RuleQuest Research of New South Wales, Australia, for the extraction of individual residential dwellings from 1x1 m pan-

sharpened IKONOS imagery over Columbia, South Carolina (Figure 5.19). In the first stage of a two-stage classifier, they determined candidate house pixels. In the second stage, they separated regions of pixels that were likely houses from those that were not.

Tullis (2003) also utilized C5.0 in a recursive manner to study optimized classification of impervious surfaces in high spatial resolution digital aerial imagery and LIDAR (light detection and ranging) remote sensor derivatives (Figure 5.20). An important consideration was the effect of spatial aggregation, investigated at various scales using *eCognition* software (left, Figure 5.20). Because of heuristics and efficient programming, C5.0 rapidly extracts rules using example data, although the process of applying the rules to unknown cases is more computationally expensive. Automated optimization in this study not only improved classification results but also eliminated unnecessary data inputs.

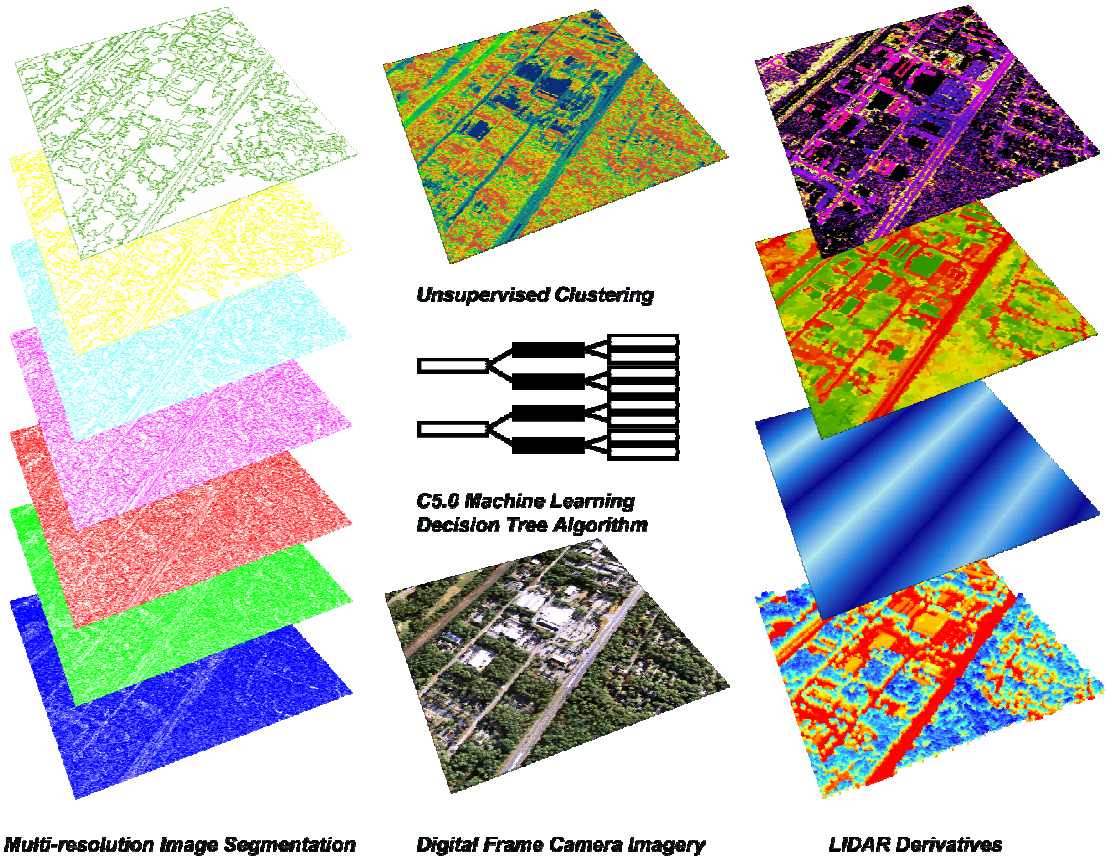


Figure 5.20. *C5.0 was used for extracting (classifying) impervious surfaces from high spatial resolution digital aerial imagery and LIDAR derivatives.*

Jensen (2005:418-419) summarizes the logic used by the C5.0 decision tree *classifier*:

1. C5.0 utilizes a recursive "divide and conquer" strategy to generate a decision tree from a training dataset.
2. The most efficient attributes are selected at each *node* of the decision tree based on the *entropy* concept from communications theory (an attribute with minimal entropy is best suited for dividing the dataset into separate groups).

3. The C5.0 generated decision tree (often difficult to read and understand) is converted to production rules (a series of if-then statements).
4. Because of the nature of definitive production rules, a major challenge is that the rules do not necessarily cease to be mutually exclusive or collectively exhaustive. *Rule confidences* and a default class are utilized in a voting procedure that ensures rule inferences can be successful in cases where they do not provide definitive answers.

The basic structure of the production rules (in either an expert system or from a machine learning decision tree) can be graphically represented by *hypotheses*, *rules*, and *conditions* (Figure 5.21). Each hypothesis is equivalent to a possible output class or potential final decision. One or more *rules*, separated by *OR* logic, each independently controls whether or not the hypothesis can be true. Each rule is comprised of one or more conditions, separated by *AND* logic, that collectively control whether or not the rule is "true" for a particular case being evaluated. In the hypothetical example of Figure 5.21, there are two hypotheses, four rules, and five variables. Arbitrary letters *x*, *y*, *p*, *l*, *z*, *n*, and *d* are used to represent individual numbers or names (e.g., 46.25, "wet"). Variables can be anything from remote sensor bands (e.g., thermal infrared 10-12 µm) to derivative structural metrics (e.g., a shape index).

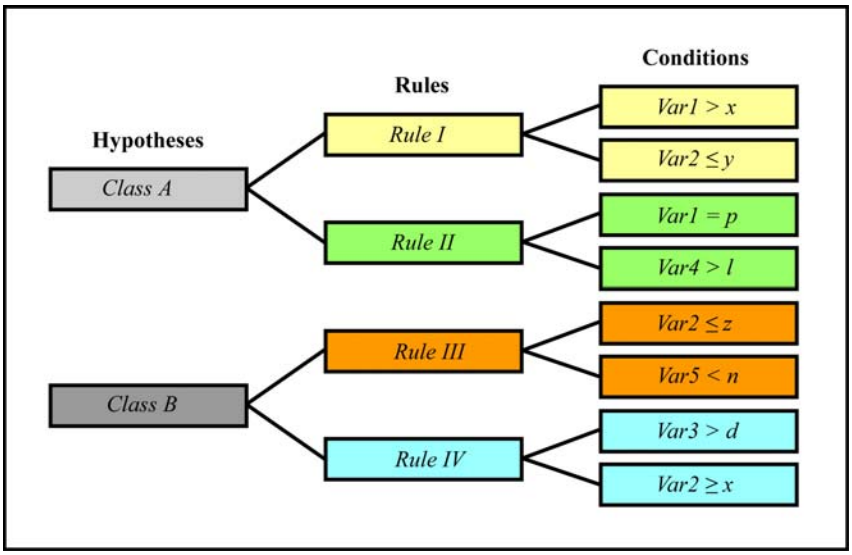


Figure 5.21. *Hypotheses, rules, and conditions form the basic structure of production rules, whether they are part of a rule-based expert system or generated using a machine learning decision tree algorithm (automated production rules).*

When unknown cases are evaluated according to specific hypotheses, rules, and conditions (referred to in expert system terminology as *inference*), the following logic applies: 1) for any rule to be evaluated as "true," all conditions connected to that rule must evaluate as "true"; 2) if any rule evaluates as "true," its associated hypothesis (e.g., output class) will *potentially* be "true."

Determination of whether a candidate hypothesis will be selected for an unknown case will depend on the nature of the specific expert system, machine learning decision tree program, or software being used for rule extraction and inference. C5.0 also

computes a confidence level for each rule extracted from the decision tree structure. These confidence levels are used in a voting procedure for selecting the output class (e.g., when two exclusive hypotheses are both candidates based on the inference of one or more of their individual rules). Furthermore, there are some instances when C5.0 cannot determine an output class using these procedures (e.g., when none of the rules in the production rule set evaluate as "true"). In this case, a *default class* is chosen. A simplistic definition of a default class is one that is represented by the greatest majority of input training cases. However, in the case of C5.0 the selection of the default class is more complex and is based on proprietary heuristics.

5.8.3. C5.0 Decision Tree Classification of Army City

The C5.0 machine learning decision tree was applied to the six co-registered geophysical datasets acquired at Army City. Predictor variables included conductivity (CON), ground penetrating radar (GPR), magnetic gradiometry (MAG), magnetic susceptibility (MS), electrical resistivity (RES), and airborne thermal infrared imagery (THERM). The pre-processed geophysical data, before and after normalization (see Section 5.2.2; Figure 5.2), were utilized in this investigation, as were previously defined training polygons (Section 5.7.3.2) for five archaeological and background classes: (1) concrete/masonry rubble, (2) pipes or large metal objects, (3) gutters, (4) other archaeological features grouped collectively, and (5) the background devoid of archaeological evidence (Figure 5.18a). As in previous supervised classifications, all classes were treated as mutually exclusive, so hard classification logic was employed here.

C5.0 operates from a simple graphical user interface and a DOS executable is also available (used here). The advantage of the latter is that C5.0 can be controlled and run in batch mode (e.g., to accomplish 10,000 runs with different inputs and parameters) and adapted to foster optimization scenarios (Tullis 2003). C5.0 works most readily on tabular data as in a vector point or polygonal feature class attribute table. This required raster-to-point-vector conversion of the data (i.e., each raster cell was converted to a vector point). Custom *ArcView 3.3* Avenue scripts have been developed to automate the process of executing C5.0 on a table of input variables, with a subset of cases labeled with a known class (the training data used in Section 5.7.3.2 and illustrated in Figure 5.18a). All 5,083 training cases were submitted to C5.0 using the customized scripts. A first script inputs training cases in a format that C5.0 can recognize while a second script extracts data from a feature class attribute table selected within the *ArcView 3.3* viewer interface. It then executes the DOS version of C5.0 based on chosen parameters. After examining the output, the user can run additional scripts that convert the C5.0 rules themselves into an Avenue script that infers output classes for all known and unknown cases in the feature class attribute table.

5.8.3.1. C5.0 classification of standardized versus un-standardized data

Two initial classifications of Army City based on the six geophysical data sets focused on a comparison between standardized and unstandardized inputs to determine which might yield a better model. Although both data sets were subjected to identical pre-processing (Section 5.2.2), the standardized data were normalized using logarithmic and power transformations so that the mean of each image was equal to zero and the

standard deviation was equal to unity. C5.0 does not require standardization of input datasets and can accept nominal, ordinal, and interval-ratio input data. Standardized and unstandardized versions of the geophysical data were therefore examined. C5.0 always evaluates its performance on its own training data and can additionally withhold some of the known cases as test cases for a more robust evaluation. Accuracy matrices are generated by C5.0 and are included in the report of each classification run. A total of 75 percent of the known cases were used for *training* and 25 percent of the cases were reserved to *test* the classifier. Standardized and unstandardized geophysical inputs were both examined and it was found that the latter produced higher accuracies by several percent, whether evaluated on the training or test data (Table 5.9).

Table 5.9. *Accuracy matrices showing C5.0 classification performance for predicting subterranean archaeological features from standardized versus unstandardized geophysical data on the training (top) and test (bottom) samples. Different random samples were employed for each run, which explains variable class sample sizes.*

	TRUE CLASS (TRAINING DATA)											
	Standardized Data						**Unstandardized Data**					
PREDICTED CLASS	**(a)**	**(b)**	**(c)**	**(d)**	**(e)**	**TOTAL/ %**	**(a)**	**(b)**	**(c)**	**(d)**	**(e)**	**TOTAL/ %**
(a) Background	**883**	0	1	6	16	**906**/97.5	**859**	0	3	4	23	**889**/96.6
(b) Pipes	5	**244**	19	16	36	**320**/76.3	1	**269**	8	1	31	**310**/86.8
(c) Gutters	7	4	**235**	14	93	**353**/66.6	13	3	**263**	14	49	**342**/76.9
(d) Concrete	24	9	12	**955**	55	**1055**/90.5	8	3	5	**1037**	28	**1081**/95.9
(e) Other	201	24	36	58	**859**	**1178**/72.9	95	7	14	40	**1034**	**1190**/86.9
TOTAL	**1120**	**281**	**303**	**1049**	**1059**	**3812**	**976**	**282**	**293**	**1096**	**1165**	**3812**
% correct	78.8	86.8	77.6	91.0	81.1		88.0	95.3	89.8	94.6	88.8	
	Overall accuracy: 83.3%						**Overall accuracy: 90.8%**					

	TRUE CLASS (TEST DATA)											
	Standardized Data						**Unstandardized Data**					
PREDICTED CLASS	**(a)**	**(b)**	**(c)**	**(d)**	**(e)**	**TOTAL/ %**	**(a)**	**(b)**	**(c)**	**(d)**	**(e)**	**TOTAL/ %**
(a) Background	**279**	0	0	2	15	**296**/94.3	**293**	0	0	1	19	**313**/93.6
(b) Pipes	2	**75**	5	5	15	**102**/73.5	4	**69**	11	4	24	**112**/62.0
(c) Gutters	3	1	**55**	7	30	**96**/57.3	3	3	**67**	4	30	**107**/62.6
(d) Concrete	8	3	11	**331**	22	**375**/88.3	4	8	6	**313**	18	**349**/89.7
(e) Other	90	16	28	36	**232**	**402**/57.7	55	8	19	18	**290**	**390**/74.4
TOTAL	**382**	**95**	**99**	**381**	**314**	**1271**	**359**	**88**	**103**	**340**	**381**	**1271**
% correct	73.0	78.9	55.6	86.9	73.9		81.6	78.4	65.0	92.1	76.1	
	Overall accuracy: 76.5%						**Overall accuracy: 81.2%**					

5.8.3.2. C5.0 results at Army City

In the unstandardized analysis the C5.0 machine learning decision tree classifier extracted a total of 138 rules from the full 5,083 training cases (the combined samples of Table 4.8). In addition to output classes a relative confidence level is also generated for each rule (Figure 5.22). The basic structure of 15 of those rules, each separated by "OR" logic, is shown in Figure 5.22, with Rule 16 (pipes/large metal) and Rule 41 (gutters) shown in detail. Each rule is composed of an if-then statement of variable size. Within a given rule, multiple conditions are separated by "AND" logic. Consequently, each

condition in the rule's if-then statement must be satisfied for that rule to be valid. Conflicting rules that are both potentially valid are voted upon according to their rule confidences. When no decision can be made, a default class is selected via a proprietary heuristic.

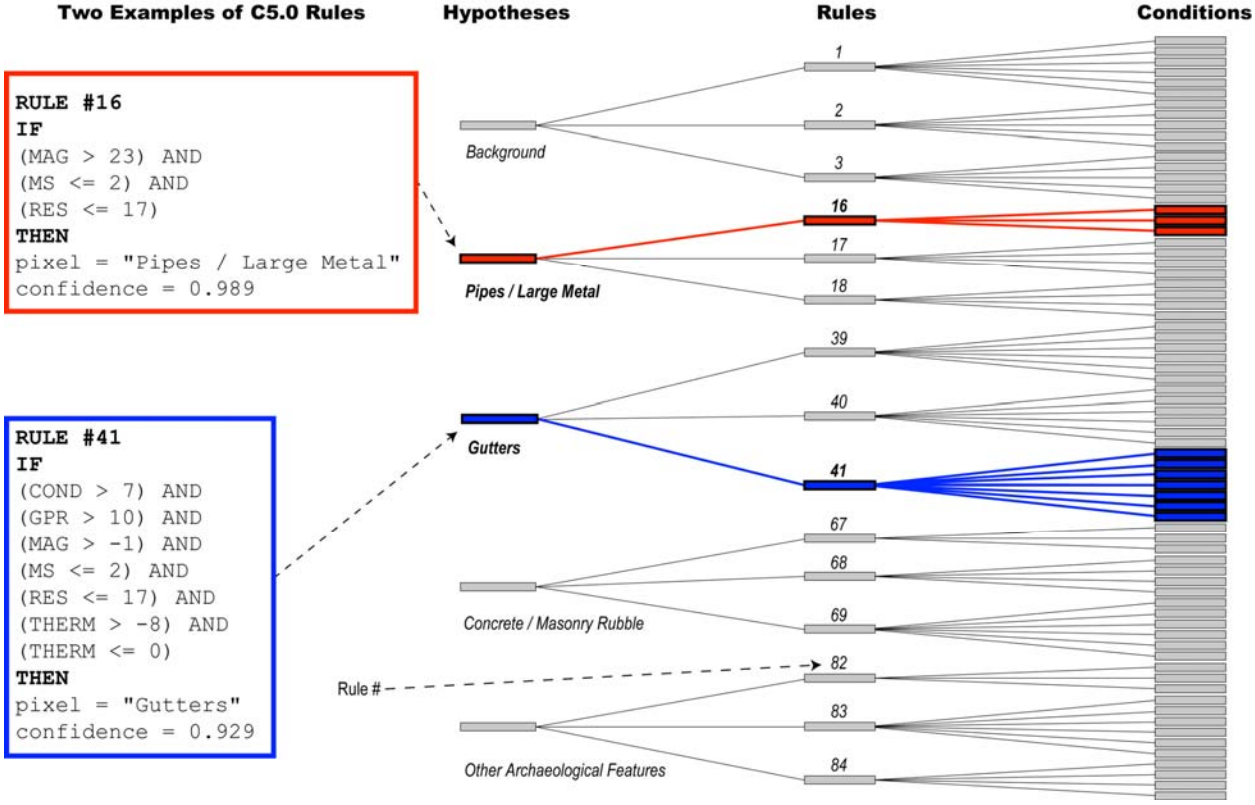


Figure 5.22. *Example of 15 rules and the rule structure for a C5.0 classification of five archaeological feature classes (background, pipes/large metal, gutters, concrete/masonry rubble, and other archaeological features) based on six unstandardized geophysical variables at Army City. Rule #16 (pipes / large metal) and rule #41 (gutters) are shown in detail. (COND=conductivity, GPR=ground-penetrating radar, MAG=magnetic gradiometry, MS=magnetic susceptibility, RES=electrical resistivity, THERM= thermal infrared.)*

Once final rules were created, additional ArcView 3.3 Avenue scripts read the rules and converted them into Avenue code. This code was then applied to *all cases*, producing classified output (Figure 5.23) where each pixel was assigned to one of the five archaeological classes (background, pipes/large metal, gutters, concrete/masonry rubble, other archaeological features) following rules diagrammed in Figure 5.22. A summary confidence level was also generated for each pixel in the study area. These are relative confidence levels that indicate how frequently a rule supports an output class. They range from "quite frequent" to "quite unusual." Pixels with high confidence levels had multiple rules supporting their output class (Figure 5.23).

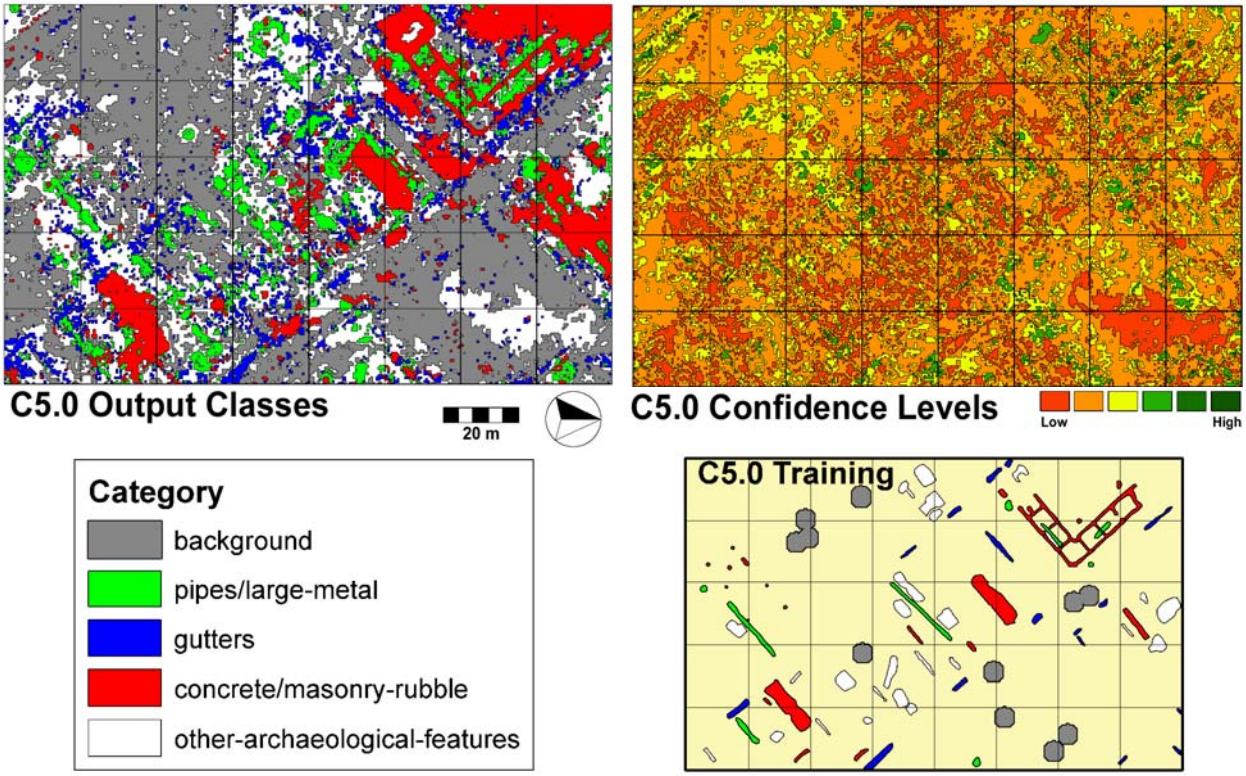


Figure 5.23. *C5.0 classification output at Army City site based on six unstandardized geophysical variables (left) with a summary confidence level generated for each pixel (right). The training polygons (lower right) represent known instances of archaeological features.*

5.8.3.3. Winnowing, cross-validation, and boosting

"Winnowing," a term that brings to mind the separation of chaff from wheat, is an excellent name for the capability C5.0 has for removing undesirable input variables from the classification process. C5.0 can pre-selectively remove variables of low importance prior to classification and estimate the relative importance of remaining variables when the winnowing option is invoked. These estimates are based on the relative increases in error when the variables are removed (or included) in the classification process. It is only intended as a rough estimate and is limited by the sample of training data selected for the classifier. The winnowing option was invoked for the classifier described above (Figure 5.22-5.23), but no variables were winnowed. This indicates that C5.0 found all geophysical variables to be of significant value when predicting the five classes of archaeological features.

Cross-validation is a useful way to evaluate the performance of a classifier when training data are at a premium (which it usually is!). Sometimes it is so expensive to collect training data that providing enough completely independent test data is problematic. Cross-validation, or the withholding of smaller numbers of cases in a recursive manner through *x* "folds" of the dataset, is a more robust way to evaluate the performance of a classifier than an evaluation on the training data itself (Table 5.9). A total of 10 folds was employed, each representing a random selection of the total with similar sizes and class distributions. For each fold a classifier was constructed using all *other* folds and tested on the "hold-out" cases of that fold. Accuracy measures for the 10

folds were then averaged to get a better idea of how the classifier will behave on additional unknown cases. Interestingly, the cross-validation resulted in an estimate of 81.9 percent accuracy for the classifier tested above. This estimate was very close to the 81.2 percent accuracy estimated by randomly withholding a subset of the known cases prior to classification (Table 5.9).

C5.0 can also implement an option called "boosting." Several classifiers are generated from a single set of training data and their combined vote is then employed when predicting the class for an unknown case. This is analogous to generating a more complex neural network for increasing accuracy. The idea is that a single classifier will make mistakes on certain cases, and subsequent classifiers can pay more attention to those mistaken areas, attempting to find rules that can be used effectively in those cases. Any single generated classifier may not be very accurate, but the combined voting power that results when invoking the boosting option often increases accuracy by a small amount. Ninety-nine classifiers were generated in this manner on the unstandardized data set used above and their individual accuracies were estimated on the training data alone, which ranged from 80-90 percent. Their combined voting power led to an aggregate classifier with an accuracy of 98 percent, however. An output map using the boosting option was not generated for practical reasons, however (significantly additional programming would be required for interfacing the 99 classifiers with the raster data).

5.8.4. C5.0 and Image Segmentation

Geographers and other scientists dealing with spatially distributed data have long been aware of the problem (or opportunity) inherent in the scale of spatial aggregation (Levin 1992). Multi-resolution image segmentation (Jensen 2005) offers the capability of spatially aggregating homogenous pixels that are adjacent to each other in a kind of multi-level hierarchical partitioning. It has been particularly noted (e.g., Hodgson et al. 2003; Tullis 2003) that classification accuracy may improve when high spatial resolution data is first segmented and the classifier then operates on a *per-segment* rather than a *per-pixel* basis.

Multi-resolution and multidimensional image segmentation was applied to the Army City geophysical data. The default options of *eCognition 4.1* include a composite of homogeneity, based on spectral (numeric) similarity between adjacent pixels (or aggregate object segments), and similarly based on shape (defined by "smoothness" and "compactness"). Spectral similarity is weighted at 90 percent and shape similarity at 10 percent. Coloring within each segment (the polygonal area or aggregate region of pixels) of each image represents the multidimensional mean pixel value (i.e., mean conductivity, mean GPR, etc.). Segmentation methods therefore yield data integrations. *eCognition's* default options were employed with arbitrary scales of spatial aggregation (scales of 2, 4, 6, 8, 10, 12, and 14). This resulted in a hierarchically partitioned series of image segments (Figure 5.24). The upper left image of Figure 5.24 represents the original pixels as a color composite (arbitrarily colored with red=resistivity, green=conductivity, blue=magnetic susceptibility). Each succeeding image shows a higher scale of spatial aggregation where each segment contains one or more polygons (or pixels) from a lower scale of aggregation.

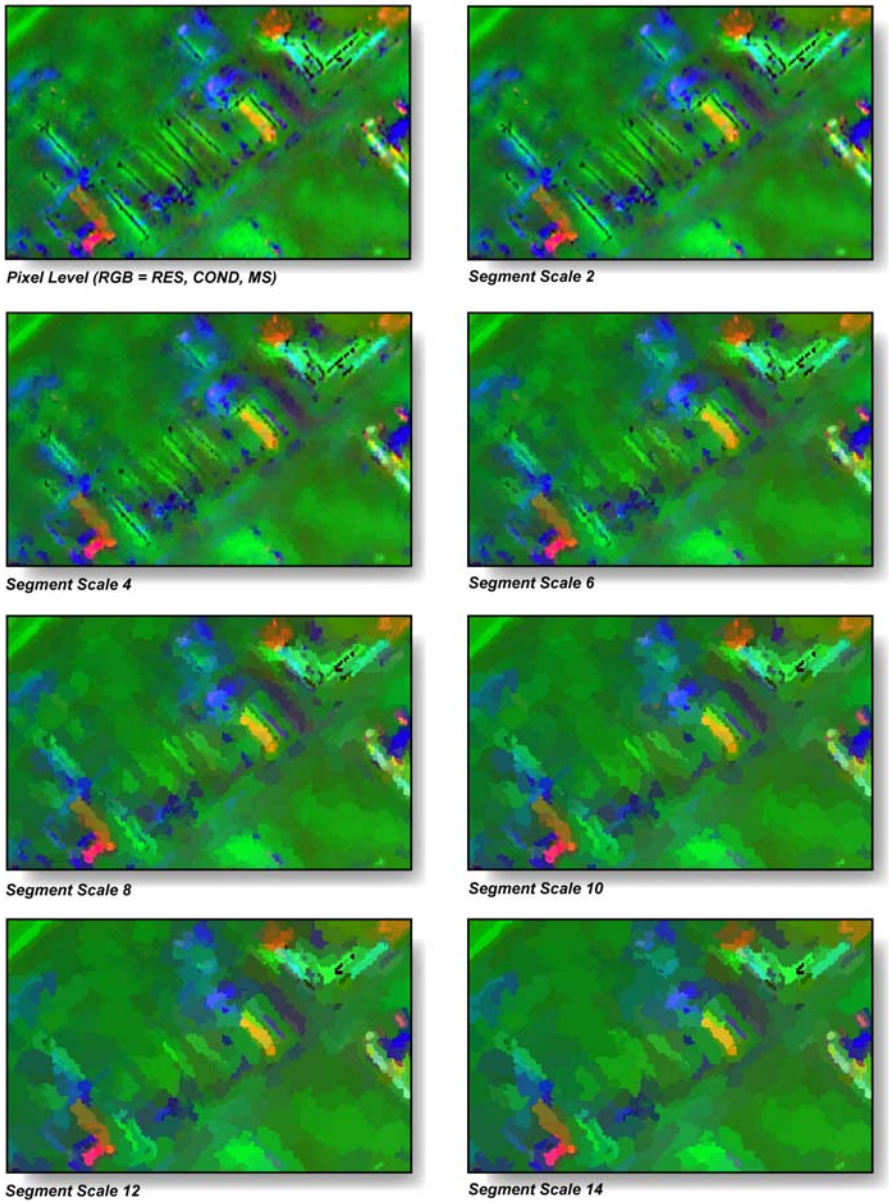


Figure 5.24. *Original scale and seven scales of multi-resolution image segmentation of Army City's geophysical data.*

As a demonstration of the potential of incorporating a spatial component into the classification process through image segmentation, one segmented image (Scale 2 in Figure 5.24) was employed to generate a new classification. The training polygons of Figure 5.18a were first intersected with that layer to provide a table of training data. This brings up an important challenge when utilizing image segmentation in classification. Unless the training data is selected on a per-segment basis, the boundaries between training polygons and image segment polygons will not match. This was the case in the present experiment where the training polygons were digitized independently based on visual appraisal of anomaly and excavation maps prior to subsequent image segmentations. This lack of correspondence caused accuracies to be much lower by approximately 10 percent (81.7 percent accuracy on evaluation of training data). In

addition, the visual output is drastically different (Figure 5.25). Besides being over-generalized, and lacking in specificity, the "gutters" class, for example, is grossly over-indicated and entire regions appear to change class membership. This initial demonstration, while by no means an exhaustive treatment of spatial aggregation in classifying archaeological data, suggests caution. One possible tactic for improvement would be to use image segments that are homogeneous based on spectral and shape criteria as training polygons. This would likely favor better results when running the per-segment classification process and may even improve the efficiency of training data selection by the user. This would be analogous to the use of a "seed tool" in standard image processing packages.

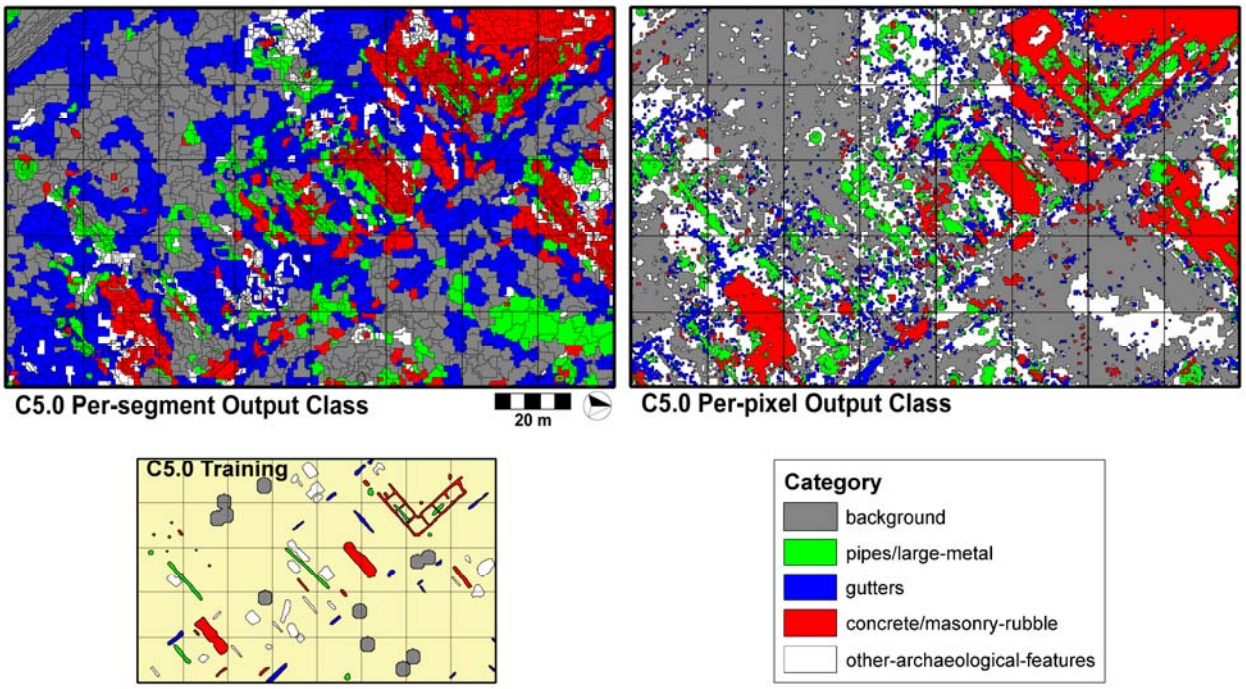


Figure 5.25. *Classification of image segments (left) versus image pixels (right, from Figure 5.23). Note that the polygon boundaries of the image segments are shown along with the training data polygons that were intersected with the image segments.*

5.8.5. Decision Tree Regressions using Cubist

Classification is a powerful concept and one that C5.0 and other algorithms do very well. This is completely different from the prediction of a numeric dependent variable (regression). C5.0 cannot be used, for example, to provide numeric estimates of biophysical variables (e.g., soy bean biomass) in remote sensing. A companion program known as *Cubist*, also developed by RuleQuest Research (New South Wales, Australia), is designed to create "regression trees," analogous to the decision trees of the preceding sections. Cubist generates rules just like C5.0, but instead of a rule inferring a class it infers a multiple linear regression model. Each rule has the following form: "If *conditions* then *linear model*" (note that "conditions" is typically plural), as in Figure 5.26. Within a rule, Cubist lists variables by decreasing relevance, so the magnetic gradiometry data are more important than resistivity in the rule illustrated in Figure 5.26, derived from the analysis described below. If all the conditions of a rule are satisfied for a given pixel, then

the associated regression model is deemed an appropriate equation for predicting the pixel's value. There may be multiple rules with conditions that are all satisfied and thus multiple rules that could be used to predict a value. With C5.0 a voting process is used to ensure that a decision can be made when multiple rules apply but point toward conflicting *classes*. Cubist, on the other hand, *averages* the values predicted by all appropriate rules (all the rules that have all their conditions met) to arrive at a final estimate.

To examine the potential of Cubist for integrating geophysical data at Army City, a tactic similar to that of Section 5.7.3.1 is employed. In that section, binary logistic regression was used to model robust anomalies indicated by two or more geophysical techniques, itself a form of integrated data (Figure 5.17a). This produced a continuous composite of all geophysical inputs that contained information richer than any single data set or targeted dependent variable (Figure 5.17b). Cubist is investigated here to model the s*um of robust binary anomalies*, a composite derived by adding together binary representations of strongest anomalies in each of the six geophysical data sets (see Figure 5.9c, Section 5.5.4). In other words, this layer illustrates how many sensors indicate an anomaly at any locus, from a minimum of zero (no anomaly indicated) to a maximum of six. Higher sums mean that more sensors reveal an anomaly, lending greater confidence to the likelihood of subsurface (e.g., archaeological) features.

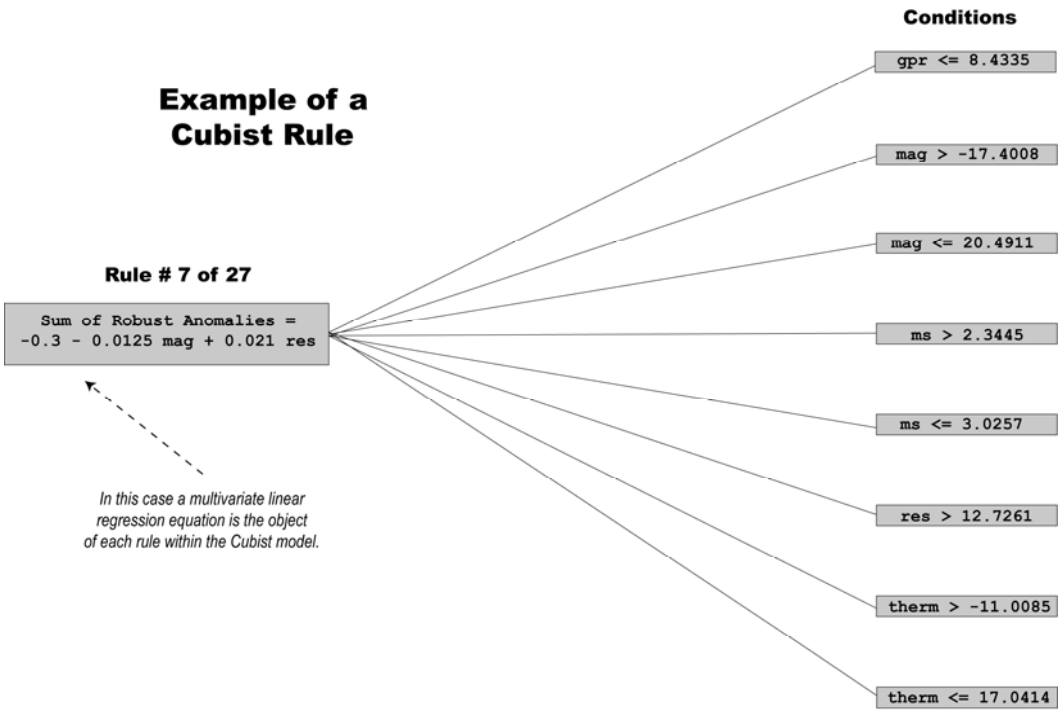


Figure 5.26. *Example of a Cubist rule developed for predicting the sum of robust anomalies. This is only one of 27 rules generated by Cubist.*

A raster layer containing the actual sum of binary anomalies (Figure 5.9c; re-illustrated in Figure 5.27) provided abundant training pixels. Half (32,000 cases) were withheld from Cubist as test cases and the other half were used for training (see Section 5.7.3.1 for a similar tactic). The results from the Cubist demonstration are impressive (Figure 5.27). Its performance is nearly identical between training and test data sets

(owing to their large sizes). Evaluation of actual versus predicted values revealed a mean absolute error of only 0.3, relative error (average error magnitude divided by the deviation from the mean value) was 0.37, and the correlation was $r = 0.86$. Computation time was longer for the regression tree, however (50.5 seconds as opposed to 0.4 seconds for the C5.0 example), but this can largely be accounted for by the fact that so many more known cases were available for analysis (64,000 as opposed to 5,083). That the model of Figure 5.27 appears nearly identical to the targeted dependent variable testifies to the power and ability of Cubist to define complex rules and relationships within a body of multivariate data. Those 27 rules, some of which are illustrated in Figure 5.26, yield hints about interrelationships between anomalies and geophysical responses while its mapping represents a continuous extension of the discrete anomaly distribution it models.

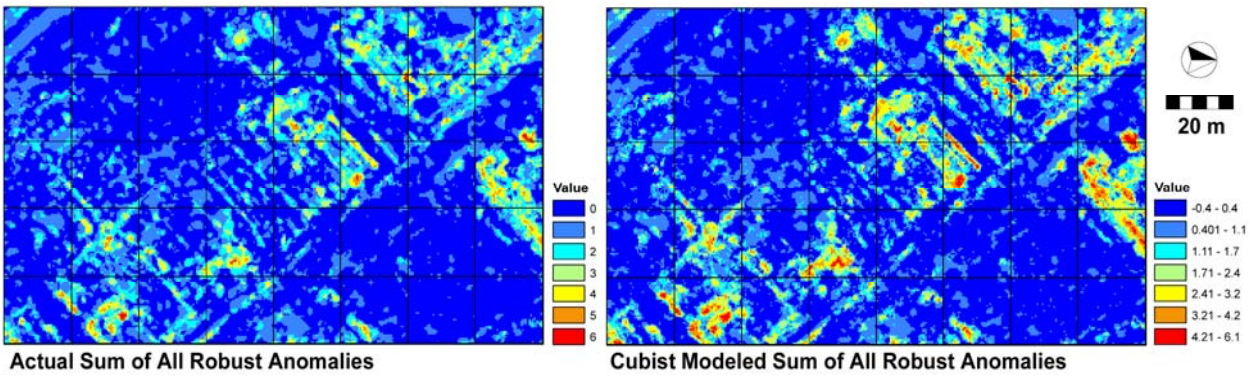


Figure 5.27. *Comparison between the actual sum of all robust binary anomalies (left) and the cubist-modeled estimates (right).*

5.9 AN EXPERT SYSTEMS APPLICATION

Kenneth L. Kvamme, University of Arkansas

Knowledge-based expert systems are used extensively in remote sensing and other applications areas. They are based on knowledge and rules extracted from human experts. This derived knowledge is converted into a system of rule-based reasoning that can be handled by a computer (Jensen 2005:408). In medicine, early expert systems have aided physicians in diagnosing and treating infectious diseases. Recent applications span many fields and include the definition of spatial objects in imagery combined with texture indices to inventory urban land cover (Moller-Jensen 1997).

A general expert system capable of distinguishing archaeological elements in a body of multidimensional geophysical data requires knowledge of many domains. An understanding of soils, geology, variations in climate, geophysical theory for each sensing device (magnetics, resistivity, radar, etc.), archaeological feature types, and human behaviors that alter archaeological deposits are some of these areas. Information in each must be converted into knowledge that is useful for making decisions about what might lie beneath the ground. Considering the wide variation of the archaeological record (which, after all, reflects the totality of all constructions by humanity throughout all places and all times), and the many kinds of soils, geologic, and environmental conditions in which they occur, such a task is probably beyond current capabilities.

Nevertheless, an attempt can be made for the specific circumstances of Army City, since a single site of short occupation length limits the range of environmental and archaeological variability that must be considered. This begs the question of how expert knowledge is converted into rules and conditions necessary for a computer to address specific problems. Knowledge is ultimately encoded within theory, rules of thumb, imagery, charts, tables, databases, and other sources. Interrogation of this knowledge base must define high-quality *rules* capable of solving the problem at issue. Such rules are expressed in a series of "IF *condition* THEN result" statements, where the condition usually represents a fact, such as *resistivity* > 50 ohm-m (Jensen 2005:410). The best way to conceptualize these rules is through a decision tree (e.g., Figure 5.21) where rules and necessary conditions for testing hypotheses are clearly charted out. The *hierarchical decision tree*, composed of hypotheses, rules, and conditions facilitates understanding of relationships—it takes an object (e.g., a pixel) described by its attributes and returns a decision. In this process, one or more hypotheses or problems are identified. Expert rules are then defined that identify one or more specific conditions that agree with the hypotheses with which they are associated (the rules may or may not be accurate, which ultimately determines performance). An inference engine—for example, a logic calculator within a GIS—then interprets the rules within the knowledge base to form conclusions. With spatial data, conclusions typically can be represented in map form.

5.9.1 Expert Rules for Army City

Following from previous sections an expert classification is pursued at Army City for the four principal archaeological classes of concrete/masonry, pipes, gutters, and "other" archaeological features, plus the background class of "no-archaeology." Each is treated as a hypothesis in an expert model (Figure 5.28). The rules associated with each vary in complexity and in the number of variables (geophysical inputs) required. All are based in geophysical theory and in expert experience with the responses of particular sensors against archaeological features in a large number of sites throughout the Great Plains and elsewhere (e.g., Kvamme 2000, 2003a, 2003b). Although quite distinct in outcome, mutually exclusive classes were only achieved by a precedence ranking of the classes.

As realized in previous classifications and general data fusion results, concrete-masonry (floors, walls, rubble piles) represents a fairly distinct class of subsurface features that are relatively easy to define. Concrete-masonry is characterized by high resistivity *and* moderate-to-high magnetic susceptibility, because these materials are highly resistant and some of the walls and floors in the site contain iron rebar or were burned, which raised the susceptibility of nearby soils (specific conditions are defined in Figure 5.28). It is convenient to next define the background class, or all loci lacking geophysical anomalies, because they too are distinct compared to the other classes. This is accomplished by requiring low or moderate values of resistivity, GPR, absolute gradiometry, and susceptibility, moderate-to-high values of conductivity, and intermediate values of thermal infrared—portions of the respective statistical distributions that did *not* point to cultural anomalies (specific conditions are defined in Figure 5.28). The remaining classes are generally less distinctive. The street gutters class is problematic because few of the associated anomalies are robust (unlike the former classes) and, indeed, all previous classifiers performed poorly against this class. On the

positive side, nearly every geophysical data set expresses anomalies associated with street gutters, although subtly (Figure 5.2). Consequently, with anomalies of moderate value, gutters require mid-range resistivity, susceptibility, absolute magnetic gradiometry, and thermal, moderate-to-large GPR, and low-to-medium conductivity (conditions are defined in Figure 5.28). Sewer and water pipes beneath the village are easily associated with very large absolute magnetic gradiometry measurements, moderate or negative values of EM conductivity, and moderate to high values of magnetic susceptibility. The first and third arise because many of the pipes are of iron or fired ceramics (which create large magnetic anomalies), while the latter occurs because metals are extremely conductive, saturate the EM signal, and cause all but high measurements (conditions are defined in Figure 5.28).

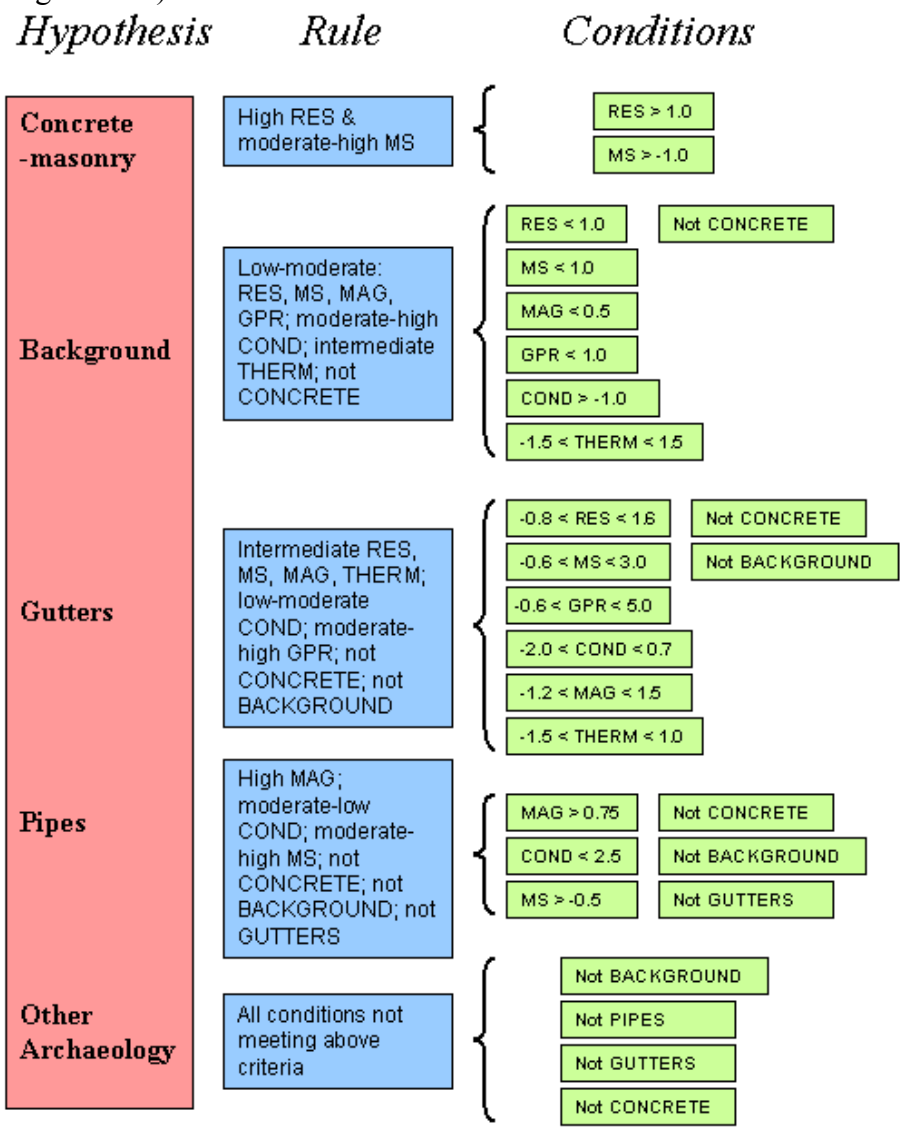


Figure 5.28. *Expert model specifying geophysical rules and conditions for identification of 5 hypotheses (archaeological classes) about subsurface features at Army City. Note that all geophysical variables are standardized.*

Finally, the last class represents the hodge-podge of poorly defined "other" archaeological anomalies. Not a homogeneous class, it has consequently caused challenges to all previous algorithms. Logically, by its relationship with the previous classes, it may be defined by them: viz. *not* background *and not* concrete/masonry *and not* pipes *and not* gutters (Figure 5.28).

The expert classification model is mapped in Figure 5.29. The result appears very robust and compares well with previous classifications, and performs better than some. This is illustrated in Table 5.10, which presents classification accuracies against the training sites defining known archaeological classes by excavation (Figure 5.18a). As anticipated, "concrete-masonry" and the "background" class are most distinctive, with highest accuracy rates. The poor performance of the "other anomalies" class is to be expected; spatially it lies adjacent to the other defined classes (Figure 5.29) and therefore contains many characteristics in common with them.

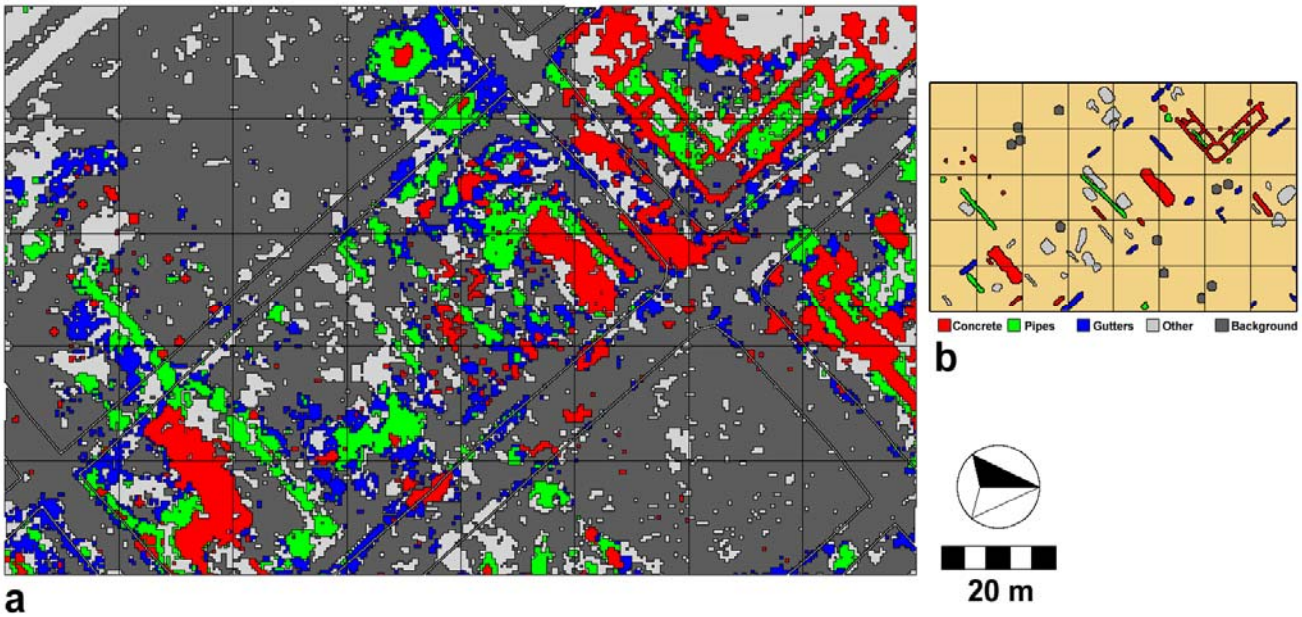


Figure 5.29. *Expert system model applied to Army City geophysical data for identification of 5 classes of subsurface features. Indicated in a): concrete-masonry (red), pipes (green), gutters (blue), other archaeological features (light gray), no archaeological features (dark gray), with loci of historic roads and alleys superimposed; b) shows original training sites.*

Table 5.10. *Classification accuracy of expert model applied to training site data.*

	ACTUAL CLASS						
PREDICTED CLASS	**Background**	**Pipes**	**Gutters**	**Concrete**	**Other**	**TOTAL**	**% correct**
Background	**468**	10	26	47	481	**1032**	45.3
Pipes	0	**292**	26	22	209	**549**	53.2
Gutters	2	37	**279**	21	301	**640**	43.6
Concrete	0	14	38	**1193**	234	**1479**	80.7
Other	36	69	80	147	**355**	**687**	51.7
TOTAL	**506**	**422**	**449**	**1430**	**1580**	**4387**	
% correct	92.5	69.2	62.1	83.4	22.5		

Overall accuracy: 59.0%; Average accuracy (columns): 65.9%

5.9.2. Discussion and Evaluation

Jensen (2005:408) notes that expert systems represent the expert's knowledge base of a subject using rules and data within a computer that can be called upon as needed. It is assumed that different problems within the same subject area can be solved using the *same* program without reprogramming. In other words, will the expert system of Figure 5.29, which is fine-tuned to the geophysical and archaeological conditions of Army City, perform well at other archaeological sites? Our assessment is *maybe*. We believe the *rules* specified in Figure 5.28 are generally correct based on expert knowledge of geophysical theory and the kinds of archaeological features targeted. Yet, the *conditions* may be too specific to the particular characteristics of Army City at the specific time the geophysical surveys were undertaken. This expert system should perform well on other archaeological sites of roughly the same type (small village), culture (American), period of occupation (early 20th century), depth (less than a half-meter), and soil type (silt-loam), but the last factor and one other introduce unpredictable variations that lends considerable uncertainty to its performance at other sites.

A particular soil type can be very idiosyncratic from place to place, varying in depth, particle size, moisture content, and according to modern vegetation cover and land use practices (e.g., plowing, fertilizing). Moreover, year-to-year climatic regimes vary considerably. The Army City geophysical data were collected in 2002, a year of severe drought in the Central Great Plains that particularly affected its soil conductivity and dielectric properties. We surmise that some of the conditions specified in the expert model of Figure 5.28 may reflect soil properties at the time and particular place of the survey, so the rules may not be optimal when applied to data collected in a wet year and somewhere else within the same soil unit, for example. Nevertheless, we believe such an expert model is a necessary start and yields the benefit of making clear the fundamental geophysical criteria related to specific archaeological phenomena (in contrast to an algorithmic approaches like C5.0 [Figure 5.22] where the number of rules is so large and their interactions so complex that they defy ready understanding). In the long term, the compilation of such rules could lead to a larger knowledge base that will increase our understanding of geophysical responses.

5.10. EVALUATION AND BENEFITS OF DATA INTEGRATION

Kenneth L. Kvamme, University of Arkansas

Integrating multiple geophysical data sets offers large potential for improved understanding of the subsurface. A single survey, for example, might reveal only part of a buried building. Integrated information from several surveys may illustrate the entire structure as well as interior components. Moreover, integrated data may *simultaneously* show relationships between conductive, resistant, magnetic, thermal, and metallic anomalies, potentially improving knowledge of features within a site, inter-sensor relationships, and enhancing overall interpretations. This possibility is hinted at by returning to computer graphics and surface rendering methods. In Figure 5.30a, b, a three-dimensional rendering of a fused and integrated geophysical data set is offered (the "binary logistic regression," which is shown below to be closely related to PCA 1 and Factor 1). This portrayal gives a fuller impression of the layout and content of Army City than the individual data of Figure 5.2, giving strong hints of the structures hidden beneath the soil (Figure 5.30c).

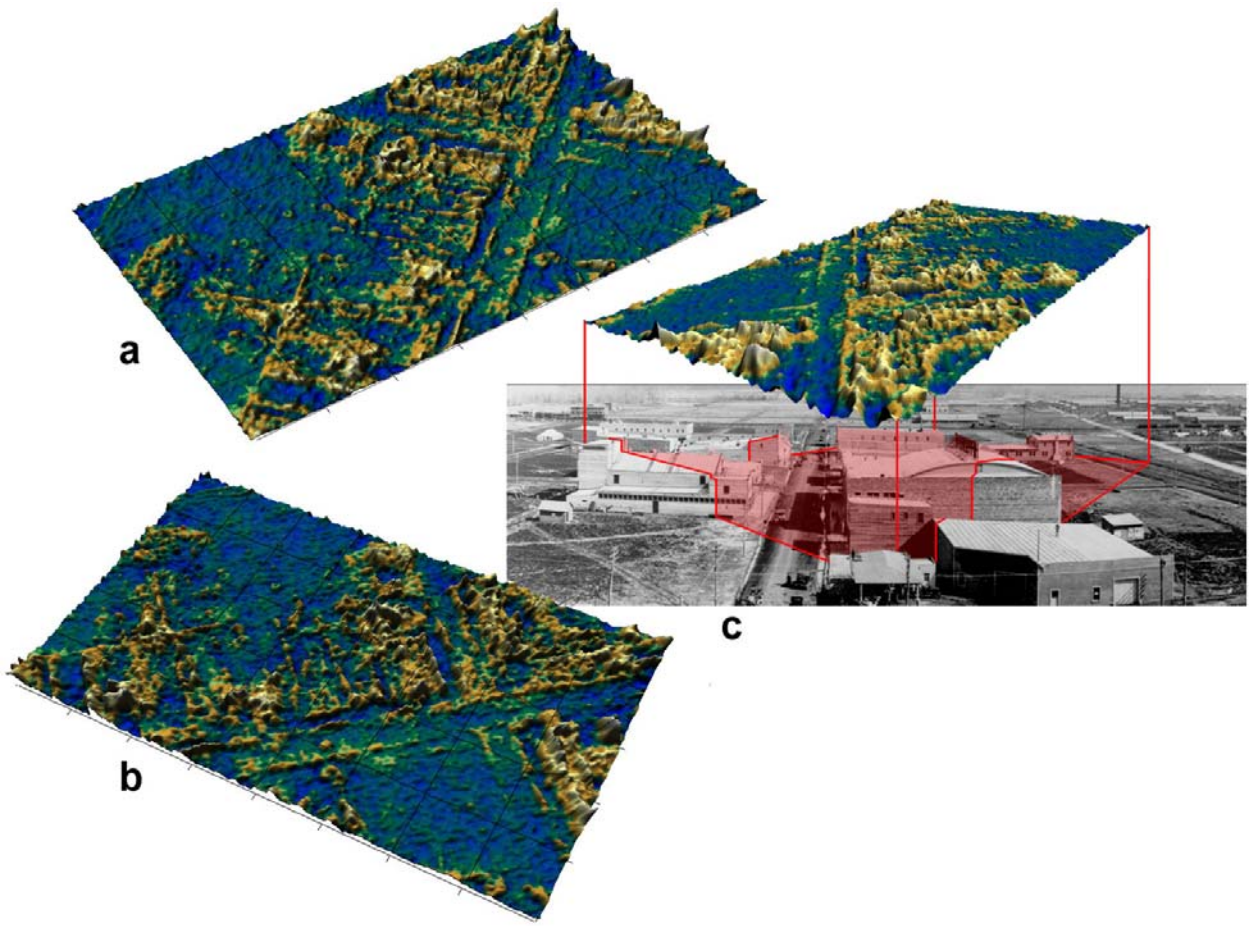


Figure 5.30. *Three-dimensional rendering methods applied to the integrated data. a) View to northwest. b) View to southwest. c) View to southeast correlating geophysical hints of ruins with historic structures in photograph.*

Graphical solutions for data integration are easy to implement and effectively combine information from disparate sources into interpretable displays. They allow complex visualizations of the subsurface, but their weakness rests in relatively low dimensionality—only 2-3 data sources may effectively be represented. Moreover, these methods are purely descriptive, yielding only images, not new data that may subsequently be analyzed. Discrete integrating methods, on the other hand, allow application of readily available Boolean operations to any number of geophysical data sets. A shortcoming is the binary maps upon which these methods are based that rely on arbitrary thresholds to define significant anomalies, while more subtle ones must be ignored. Continuous data integrations can yield insights beyond the capabilities of other methods. Robust *and* subtle anomalies may be simultaneously expressed, producing composite imagery with high information content. Interpretive data are also generated in the form of principal component scores, factor loadings, or regression weights that add to understanding of interrelationships and underlying dimensionality. Supervised and unsupervised classification methods are noteworthy because they introduce a predictive aspect to the integrating process. Patterns in these data fusions may point to anomalous conditions much less visible in any single data set that might otherwise be overlooked and ignored. They therefore offer a possible means to augment prospecting capabilities (see Johnson and Haley 2004). Although the approaches to geophysical data integration examined here span a wide range of commonly available techniques, they are by no means exhaustive. A host of other supervised and unsupervised classification algorithms exists as well as new context-based, image segmentation, and intelligent knowledge-based methods (Schowengerdt 1997; van der Sande et al. 2003).

If the foregoing results can be generalized, it is that robust anomalies exist in the data and tend to dominate any form of fusion, regardless of the method employed. The consequence is amazingly parallel results between widely different forms of integration. A correlation matrix between many of the results supports this view (Table 5.11). Correlations in the .5-.7 range are common, with many in the .8-.9 range. One relationship between very different methods is especially high. The correlation between Factor 1 and binary logistic regression is r=.985, which occurs because the factor score coefficients (Table 5.5) show an uncanny proportional resemblance to the regression coefficients (given in text). Factor 1 characterizes the principal axis of variation in the data, from anomalous through common conditions, and this is exactly what the binary logistic regression attempts to model. Given apparent levels of redundancy, it is also obvious that certain integrations are nearly independent of the others, and these tend to be the lower-order principal components and factors, which are *designed* to be independent. Consequently, they really should be considered as offering new information about subsurface variation.

Table 5.11. *Pearson correlation coefficients between selected integrating methods.* ***Bold****: r ≥ .85; Italics: r ≤ .20.*

INTEGRATING METHOD		Discrete					Simple Math				Multivariate Statistics						
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**
1	Union	1	.80	.53	.34	.52	.42	.53	.67	.55	.52	*.19*	*.01*	.55	*.01*	.56	.51
2	Binary sum		1	.84	.44	.66	.62	.75	.44	.63	.72	*.20*	*-.01*	.75	*.07*	.77	.70
3	2 or more[1]			1	.41	.62	.57	.65	.41	.54	.65	*.16*	*-.01*	.66	*.08*	.68	.63
4	Poly sum[2]				1	.59	**.85**	.73	.48	.29	.78	-.25	*.04*	.64	.51	.65	.72
5	K-means 2[3]					1	.72	.74	.65	.52	.79	*.12*	*-.05*	.78	*.17*	.79	.76
6	Sum						1	**.90**	.63	.43	**.93**	-.23	*-.03*	.80	.54	.80	**.86**
7	Product							1	.76	.54	**.96**	-.03	*-.05*	**.89**	.37	**.91**	**.86**
8	Maximum								1	.84	.72	.21	*.01*	.75	*.06*	.75	.73
9	Variance[4]									1	.52	.36	*.13*	.62	*-.15*	.52	.63
10	PCA 1										1	*0*	*0*	**.94**	.35	**.95**	**.92**
11	PCA 2											1	*0*	.35	**-.94**	.27	*.09*
12	PCA 6												1	*0*	*0*	*-.04*	.30
13	Factor 1													1	*0*	**.98**	**.90**
14	Factor 2														1	*.08*	.24
15	Bin logistic[5]															1	**.88**
16	Mult logistic[5,6]																1

[1]Anomalies indicted by 2 or more methods; [2]Sum of 3-class polychotomies; [3]K-means 2-cluster solution; [4]Standard deviation used; [5]Raw regression function used; [6]Inverse of background class

Which integrating methods are best? That may depend on purpose. Some yield visually pleasing results that appear to well-combine available information while others may seem less revealing but may offer interpretive or predictive potential. If a goal is to define discrete classes of anomalies that may be subsequently interpreted through comparison with primary data then categorical methods may be best. If a goal is merely a continuous-tone image that represents most of what is known about the subsurface then a composite color graphic or mathematical-statistical integration may be most suitable. Of course, continuous methods yield quantitative data that may subsequently be analyzed, plus regression weights, PCA scores, or factor loadings that give additional insights beyond graphical representations, important for improved understanding of the subsurface and its interaction with geophysical methods. In practice, a variety of different integrating methods may work best in practice, because each variation may give new insights about a different aspect of the subsurface.

The potential impact of the intra-site "predictive" models explored here is unknown, but based on these results it can be argued that a clearer picture of the subsurface is realized, compared to the individual geophysical maps. Consequently, such models might potentially be used to better target precision excavations over predicted archaeological features of importance, thereby making excavations more efficient and cost-effective. By more clearly imaging the totality of information about the subsurface from all sources, a better understanding of site content, structure, and organization may also be achieved. Since excavation of entire sites or settlements, or even large areas of

them, is impossible owing to funding limitations and ethical concerns, it may be only through integrated remote sensing that real understandings of the content, structure, and extent of archaeological sites may be achieved. It is anticipated that the methods pioneered here provide an important step in the direction of that goal.

In this age of high technology, itself responsible for the geophysical data, it is apparent that approaches to data presentation and integration have generally been neglected. The computer graphic, GIS, mathematical, statistical, and knowledge-based solutions offered here provide a glimpse of what is possible in today's software environments.

5.11. DATA INTEGRATIONS FROM OTHER STUDY SITES

Eileen G. Ernenwein, University of Arkansas

The data integrations carried out on the Army City data (Sections 5.3-5.9) formed a pool of methods available for application to the other project sites. It must be remembered that Army City formed an unusual data set in this project. It was (1) the most multidimensional of all project sites, with *six* geophysical data sets available that each indicated significant cultural anomalies, (2) a large contiguous area was fully surveyed by *all* methods, and (3) the only aerial data set that revealed subsurface archaeological features in the assigned study blocks occurred at this site. In the other project sites—Pueblo Escondido, Silver Bluff Plantation, and Kasita Town—only three geophysical data sets that reveal something about the subsurface are normally available and two of these sites only possess partial surveys by some of the available methods. Consequently, with the knowledge gained by the myriad methods applied at Army City, and with the poorer data available at the other sites, a more limited suite of data fusion techniques was pursued at these places. The more insightful of these results are summarized here. Interpretations of the integrated data are presented in Section 5.12.

5.11.1. Pueblo Escondido

Data fusions for Pueblo Escondido were created using the three most informative data sets: ground-penetrating radar (three slices), magnetic susceptibility, and magnetic gradiometry. Each data layer was standardized to have a mean of zero and standard deviation of one (Section 5.2.2). A cross-correlation of these layers shows that the three types are statistically independent but, as expected, the GPR slices are more strongly correlated (Table 5.12). The uncorrelated nature of the three data types is beneficial, but is also a limiting factor for certain types of data fusion (see Section 5.7.1). Multiple lines of evidence for the same feature, which can add confidence to interpretations, are also less likely to be realized. The benefit is that the three different methods offer independent sources of information, justifying the use of multiple geophysical methods in archaeological investigations. As demonstrated in Sections 5.3-5.9, an immense variety of data fusions can be created with varied and often useful outcomes. In this section, some of the more successful results for Pueblo Escondido are illustrated.

Table 5.12. *Correlation matrix between geophysics acquired at Pueblo Escondido.*

	GPR 15-31 cm	GPR 31-47 cm	GPR 47-63 cm	Magnetometry	Magnetic Susceptibility
GPR 15-31 cm	1.00	.66	.54	.05	0
GPR 31-47 cm	.66	1.00	.69	.04	.04
GPR 47-63 cm	.54	.69	1.00	.04	.05
Magnetometry	.05	.04	.04	1.00	.16
Magnetic Susceptibility	0	.04	.05	.16	1.00

5.11.1.1. Integrating GPR data

Given three moderately correlated GPR data sets, each representing a different time-slice or level beneath the surface, it is beneficial to integrate them into a *single* composite representing all GPR anomalies, in order that GPR will not over-dominate some of the later integrations with other data types. This was accomplished through a principal components analysis between the three time-slices. This PCA is very different than the one described for Army City in Section 5.7.1. That analysis was performed to integrate the multidimensional geophysical data into a reduced set of components that best summarized the total geophysical variation. Here it is employed as a data reduction technique to reduce the three GPR slices to a single dimension that represents all three. The first component is illustrated in Figure 5.31. It represents 76 percent of the GPR variance, with the remaining variance contained in the second and third components, principally illustrating unwanted noise. The first component therefore contains the majority of useful information in the three depth-slices, and represents a fusion of the GPR information. Although depth information is lost, the simultaneous portrayal of anomalies from all depths is very useful for interpretation. This result forms an input to several other fusions, below.

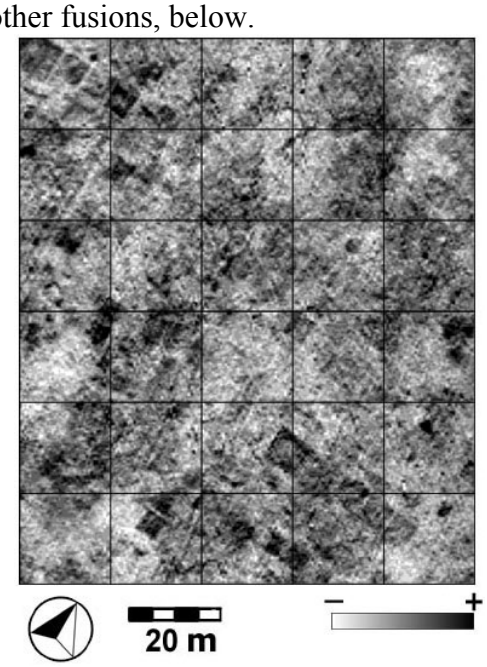


Figure 5.31. *Fusion of the three GPR time-slices from Pueblo Escondido in the first principal component of a PCA analysis.*

5.11.1.2. Integration by summing ranked polychotomies

The reclassification of geophysical data into ranked polychotomous categories as a basis for data integration was discussed in Section 5.5.6. Ranked categorical data sets were created by reclassifying the three normalized data layers from Pueblo Escondido using boundaries based on standard deviations. The following five classes were created: $(^{-}2) < {}^{-}2\sigma$, $(^{-}1)\ {}^{-}2 - {}^{-}1\sigma$, $(0)\ {}^{-}1 - {}^{+}1\sigma$, $(1)\ {}^{+}1 - {}^{+}2\sigma$, and $(2) > 2\sigma$. This result is illustrated for magnetic susceptibility, magnetic gradiometry, and the first principal component of all GPR slices in Figure 5.32a-c. Following the protocol of Section 5.5.6, these data were summed to form an effective integration (Figure 5.32d). This result shows robust positive and negative anomalies simultaneously.

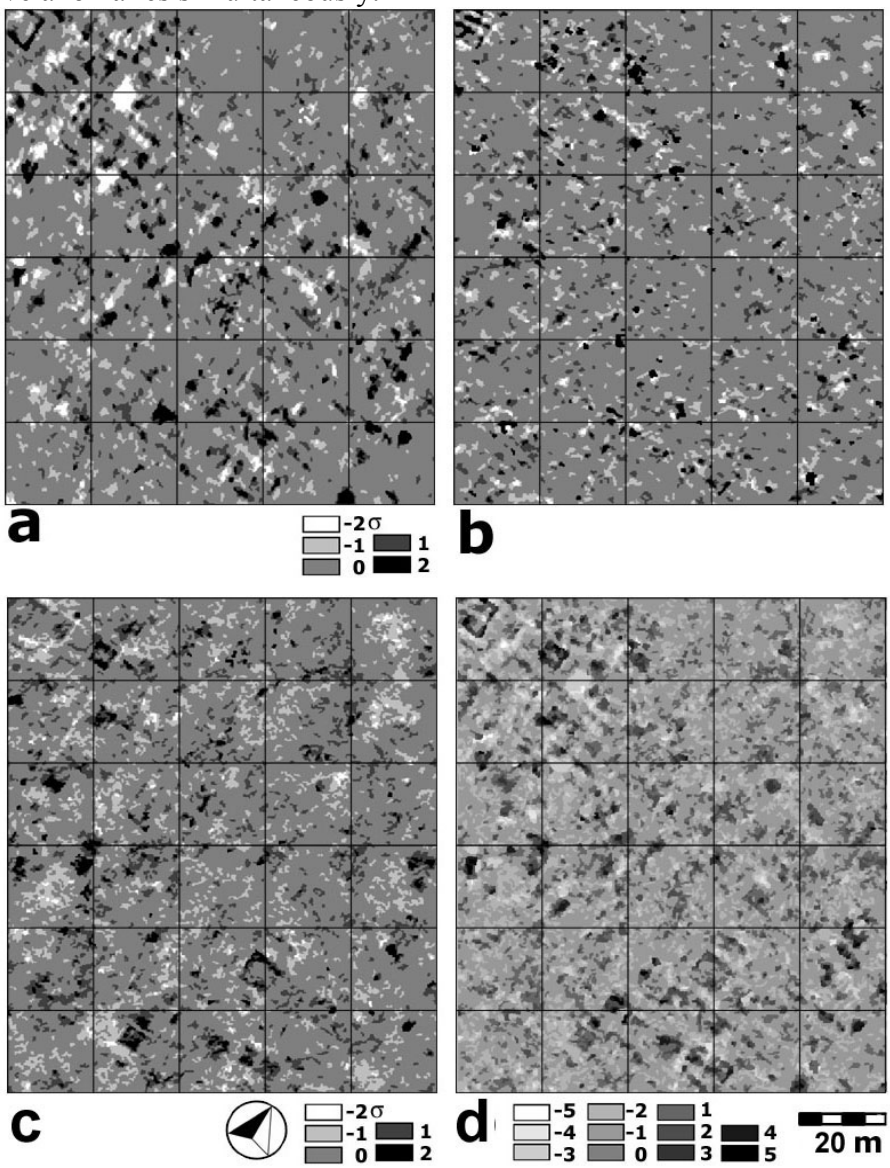


Figure 5.32. *Ranked categories based on standard deviation thresholds of a) magnetic susceptibility, b) magnetic gradiometry, and c) the first principal component of the GPR data. d) The sum of the classified inputs, a-c, representing a data fusion.*

The ranked classifications of Figure 5.32a-c may also be combined as red-green-blue (RGB) color composites (Section 5.4.2). In this example only *positive* classes 1-2 are employed to minimize clutter (Figure 5.33a). Any overlap between the data sets produces cyan, magenta, yellow, or white, but these colors occur in few areas. With a relative absence of overlapping colors, this composite clearly shows the lack of correlation between the data types (Table 5.12). At the same time, this result reveals positive anomalies for the three data types simultaneously as any colored region. Black represents locations where all data types had values less than one standard deviation above the mean (Figure 5.33a).

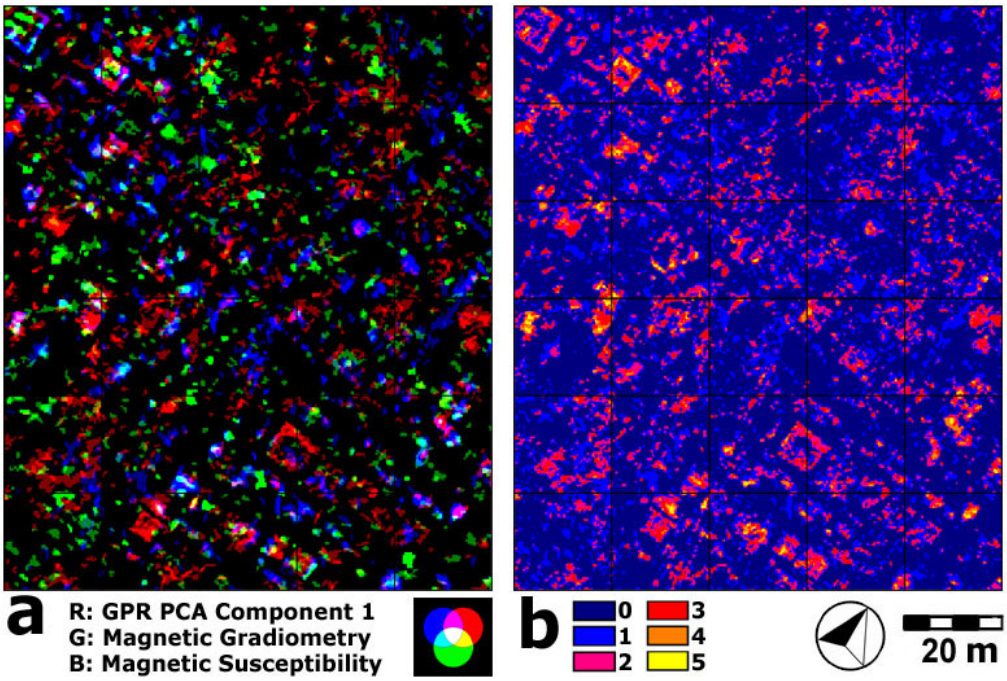


Figure 5.33. *Data fusions based on discrete categorical methods at Pueblo Escondido. a) RGB composite of positive ranked categories of GPR, magnetic gradiometry, and susceptibility. b) Sum of binary discrete classifications of GPR, magnetic gradiometry, and susceptibility, with larger sums pointing to more robustly indicated anomalies.*

5.11.1.3. Integration by binary summation

Discrete binary data fusions were also investigated. Binary mappings of most robust anomalies in magnetic susceptibility, magnetic gradiometry, and each of the three GPR slices were created using arbitrary thresholds (see Section 5.5.1). These results were then summed (Section 5.5.4). With five binary inputs the sum can range between zero (no anomalies indicated) through five (anomalies indicated in all inputs). Using the three GPR slices separately did bias the result toward GPR, but this was beneficial because it was the most revealing technology at Pueblo Escondido (Figure 5.33b). The GPR results are therefore emphasized, and areas where the other data sets overlap are easy to see (mostly orange and yellow). Although the two magnetic data sets are de-emphasized, if only one GPR input layer is used little overlap results, yielding an outcome similar to a

Boolean union of binary inputs (Section 5.5.2). Binary sums are probably most effective when using several data sets with a fair amount of anomaly overlap.

5.11.1.4. Integration by the maximum function

Several tactics for integrating data make use of the GIS *maximum* function (see Section 5.6.4), which incorporates the largest anomaly in each of the input layers on a per-pixel basis. Three are presented in Figure 5.34. In the first, the maximum value was selected among *all five* geophysical layers, including the separate GPR slices (Figure 5.34a). In the second, the first GPR principal component was used instead of the three individual slices, together with the two magnetic data sets (Figure 5.34b). Both fusions are very similar, but the latter contains less high frequency clutter and appears to better show some of the subtle details around the pithouse anomalies, especially in the southern portion of the survey area (see Section 5.12.3.2 for initial interpretations of these data). The final method maps the maximum of the three GPR time-slices in red, which is covered by a translucent blue-tinted magnetic susceptibility image and a translucent yellow-tinted magnetic gradiometry image (Figure 5.34c; see Section 5.4.3 for discussion of translucent graphics). All three results form effective integrations of the remote sensing information at this site, although with many strong parallels.

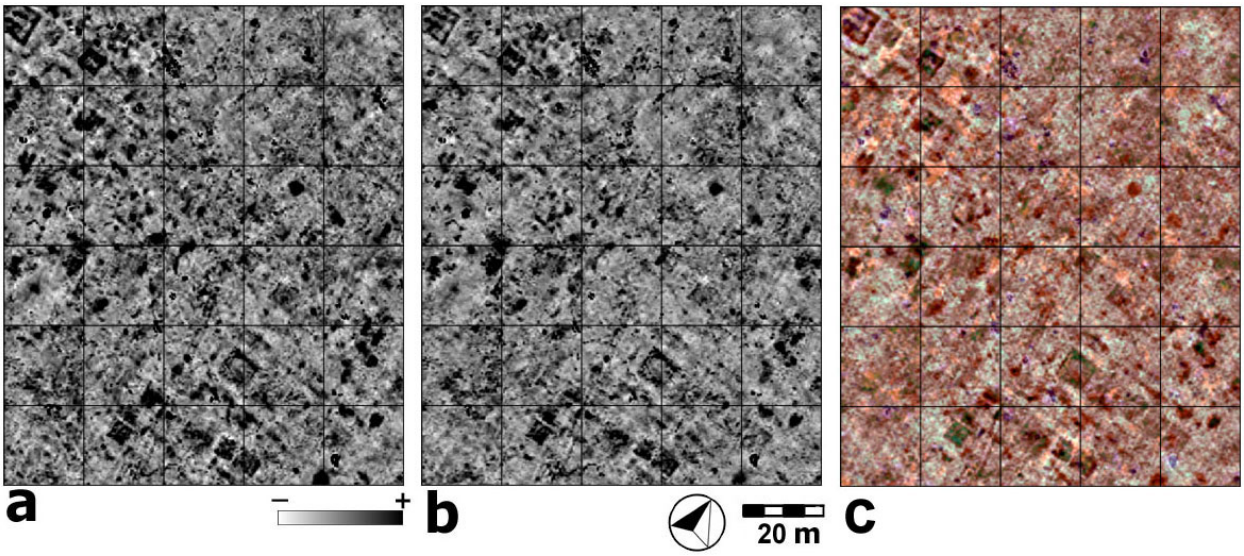


Figure 5.34. *Data integrations at Pueblo Escondido based on the GIS* maximum *function. a) The maximum of three GPR slices, magnetic susceptibility, and magnetic gradiometry. b) The maximum of the first component of the GPR PCA, magnetic susceptibility, and magnetic gradiometry. c) The maximum of three GPR slices (red), mapped against magnetic susceptibility (blue) and magnetic gradiometry (yellow) in translucent overlays.*

5.11.1.5. Integration by statistical algorithms in supervised classification

Perhaps the most sophisticated fusion results at Pueblo Escondido were achieved using the statistical methods of logistic regression and discriminant analysis. For both methods, training sites were selected from robust anomalies in the geophysical data that could be confidently interpreted (these investigations were made prior to the field evaluation and testing phase of the project where positive identifications were assigned to

many anomalies; see Section 8.14). Training sites for the logistic regression were selected from portions of robust anomalies in each data set that expressed geometric shapes suggestive of a cultural origin. The total number of training samples (pixels) was $n = 307$ (out of $N = 192,000$), making up only .2 percent of the entire survey area.

The five geophysical data layers were then employed to generate a probability surface for the presence of these anomalies, as outlined in Section 5.7.3.1. A logistic regression analysis yielded the following function:

$$L = -7.686 + .398*GPR9\text{-}16cm + .002*GPR15\text{-}31cm + .960*GPR31\text{-}47cm + .099*GPR47\text{-}63cm + .159*MAG + .575*MS$$

associated with a pseudo-R^2=.23. Since the input layers were standardized the coefficients can be meaningfully compared. Interestingly, the GPR slices received widely varying weights, ranging from .002 (virtually no contribution) to .960. The weights for the magnetic layers seem fitting given that magnetic susceptibility seems to show cultural patterns throughout the survey area, where magnetic gradiometry, with a coefficient of only .159, showed clear cultural anomalies in only a handful of locations. A logistic transformation, $p = (1+exp(-L))^{-1}$, of this function yielded the probability surface shown in (Figure 5.35a). This surface was then classified to indicate the distribution of most robust anomalies by the model (Figure 5.35b).

As a variation, linear discriminatory analysis was also explored using the same input layers (Klecka 1980). This method is similar to logistic regression, but unlike that nonparametric technique, it is based on a multivariate normal model. Results, however, are often closely similar. For this data integration training sites for feature presence or absence were generated through a random sample taken from a binary layer showing areas where at least *two* geophysical sensors simultaneously indicated anomaly presence (see Section 5.5.5). The resulting linear discriminatory function

$$D = -131.583 + .152*GPR9\text{-}16cm + .384*GPR15\text{-}31cm + .413*GPR31\text{-}47cm + .154*GPR47\text{-}63cm + .101*MAG + .379*MS$$

was then employed to map anomaly presence and absence based on a maximum likelihood decision rule (Figure 5.35c). The coefficients of this model show a more balanced weight between the various layers compared to the regression model, but their mappings appear very similar in form. It is obvious from this result that both methods yield closely parallel findings despite different training sites.

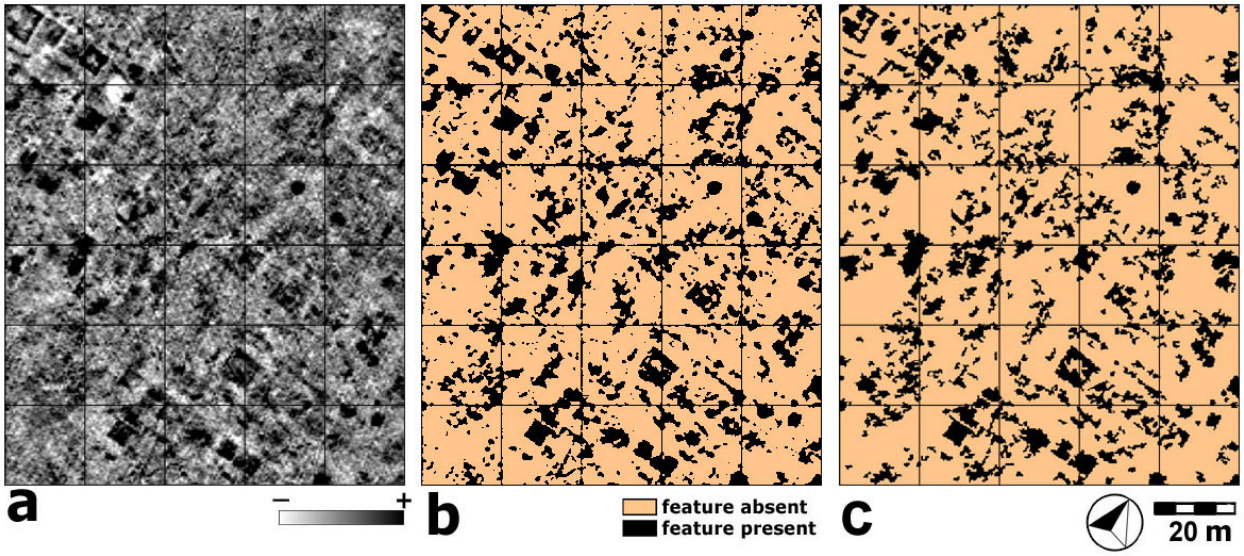


Figure 5.35. *Integration results at Pueblo Escondido by statistical algorithms based on supervised classifications. a) Logistic regression probability surface for robust anomaly presence. b) The logistic regression probability surface is classified to indicate most robust anomalies. c) Linear discriminatory function mapping of robust anomalies based on maximum likelihood criterion.*

5.11.1.6. Model comparisons

The previous findings showed broad similarities between two models derived in different ways. Indeed, a perusal of any of the preceding graphics reveals the same anomalies appearing time and again regardless of the form of data integration. This indicates that certain strong anomalies at Pueblo Escondido tend to dominate results regardless of subsequent data processing. A correlation matrix between the continuous fusion results of this section verifies that this tends to be the case with moderate to high correlations illustrated between the various methods (Table 5.13).

Table 5.13. *Correlation matrix between continuous fusions at Pueblo Escondido.*

	GPR C1	Max 1	Max 2	Max 3	Regression
GPR C1	1.00	.4734	.6446	.6489	7540
Max 1	.4734	1.00	.8325	.6850	.6347
Max 2	.6446	.8325	1.00	.7253	.6709
Max 3	.6489	.6850	.7253	1.00	.7670
Regression	.7540	.6347	.6709	.7670	1.00

(KEY: GPR C1 refers to Figure 5.31; Max 1, Max 2, and Max 3 refer Figure 5.34a-c, respectively; Regression refers to Figure 5.35a.)

5.11.2. Silver Bluff Plantation, South Carolina

Three geophysical data layers from Silver Bluff Plantation were chosen as inputs for data fusion: magnetic gradiometry, electrical resistivity, and one GPR time-slice representing 39-105 cm below the surface. Though several GPR slices were available, only one was chosen for this purpose. A principal components analysis between the several time-slices (see Section 5.11.1.1) revealed that no more than 30 percent of the

variance could be captured in any one component, so any hopes of combining several slices into a single result containing most of the information were lost. The chosen GPR slice appeared to best show cultural anomalies when compared to each principal component, however. It should be remembered that Silver Bluff Plantation was covered with hundreds of chopped up pin flag pieces strewn about the site (Section 3.3.4.3), creating an enormous geophysical clutter of unwanted anomalies that generally obscure what lies beneath the surface. Dropping magnetometry as a meaningful layer was considered, but with so few data sets to work with it was retained. Since magnetic dipolar anomalies present challenges to many data fusion algorithms—because robust anomalies illustrate positive *and* negative values (see Section 5.2.2)—absolute values were taken. In addition, electrical resistivity measurements were inverted in many data fusions because low values often corresponded with positive GPR anomalies. All data layers were processed with low- and high-pass filters to reduce noise and remove broad regional trends, respectively, and standardized so that $\mu = 0$ and $\sigma = 1$ (see Section 5.2.2). A correlation matrix between the three layers indicates they are statistically independent (Table 5.14). Such a relationship effectively precludes the use of some types of fusion methods, such as principal components analysis (Section 5.7.1), while it makes others more effective. A great variety of fusion algorithms were experimented with for these data, and a sample is presented here.

Table 5.14. *Correlation matrix between geophysical data layers at Silver Bluff Plantation.*

	GPR 39-105 cm	**Magnetic Gradiometry**	**Electrical Resistivity**
GPR 39-105 cm	1.00	.01	-.08
Magnetic Gradiometry	.01	1.00	-.07
Electrical Resistivity	-.08	-.07	1.00

5.11.2.1. Integration by RGB color composites

Due to the uncorrelated nature of the input data layers, red-green-blue (RGB) color composites (Section 5.4.2) prove very useful for summarizing the majority of information in a single image. When the three original input layers are combined in this way, the result shows the positive anomalies quite clearly as primary colors (Figure 5.36a), because there is little overlap between anomalies (Table 5.14). Increased clarity of data patterns may be obtained using a similar RGB combination applied to ranked versions of the input data layers. Since each had been previously standardized, this required only a simple reclassification into five categories for each layer: (1) $< -2\ \sigma$, (2) -2 to -1 σ, (3) -1 to +1 σ, (4) 1 to 2 σ, and (5) $>2\ \sigma$. When combined (Figure 5.36b), the result is similar to the continuous version, but the anomalies appear much clearer because of the discrete boundaries and reduced clutter. The rectangular pattern of anomaly distributions (Section 5.14.4) is emphasized in these data fusions that seem to point to a wide distribution of culturally patterned landscape.

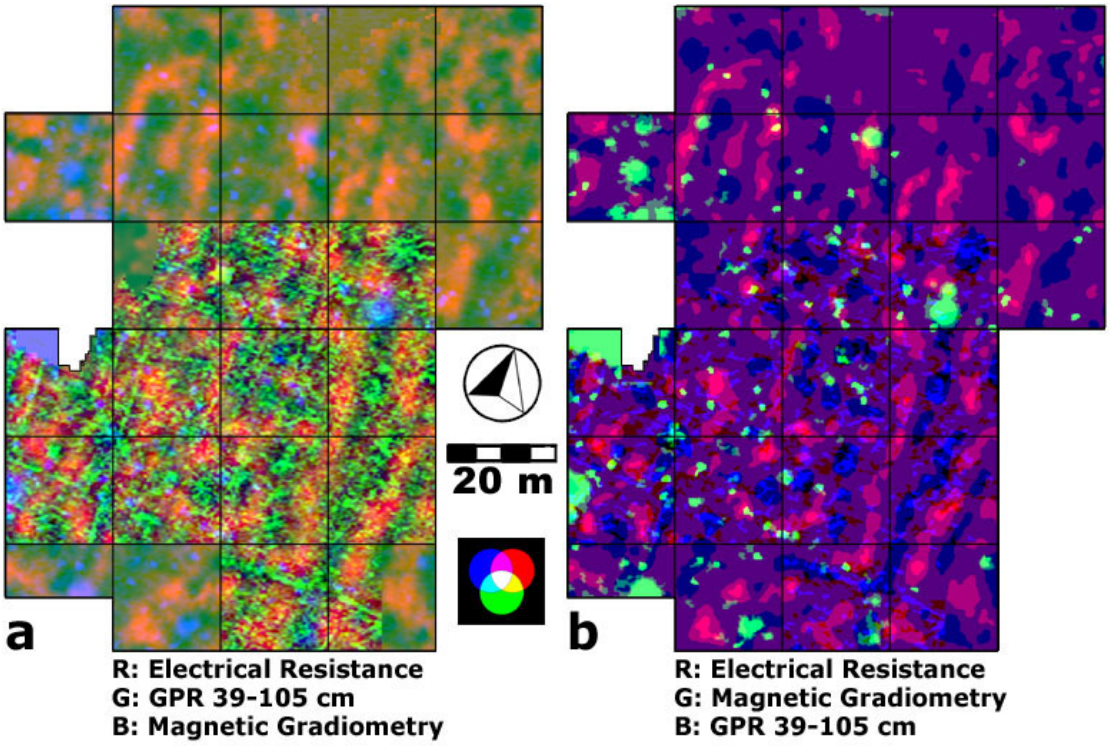


Figure 5.36. *RGB fusion results at Silver Bluff Plantation. (a) Composite of standardized continuous data layers. (b) Composite of ranked discrete versions of original inputs.*

5.11.2.2. Continuous data integrations

This investigation utilizes GPR (39-105 cm below surface), magnetic gradiometry, and the *inverse* of electrical resistivity (i.e., -1.0 x resistivity). The inversion was undertaken because positive GPR anomalies tend to coincide with negative resistivity anomalies in this data set. In addition, a constant of 100 was added to all layers to avoid multiplication of negative numbers.

Integrating geophysical layers in their continuous forms has the advantage of preserving subtlety, which can aid interpretation if different sources help to complete larger anomalies. This is evident in results showing the product of the three available layers (Section 5.6.2) and the per-pixel maximum (Section 5.6.4), illustrated in Figure 5.37. When cross-multiplying the layers to create a data product the more robust anomalies are accentuated. Similarly, when the maximum value for each pixel is extracted, the final outcome emphasizes robust positive anomalies. Unlike the data product, the per-pixel maximum tends to accentuate positive anomalies, while negative anomalies become part of the background. This result appears somewhat more interpretable than the data product. Both of these fusions emphasize numerous magnetic "point" anomalies, most of which probably derive from the general distribution of pin flag wires covering the surface.

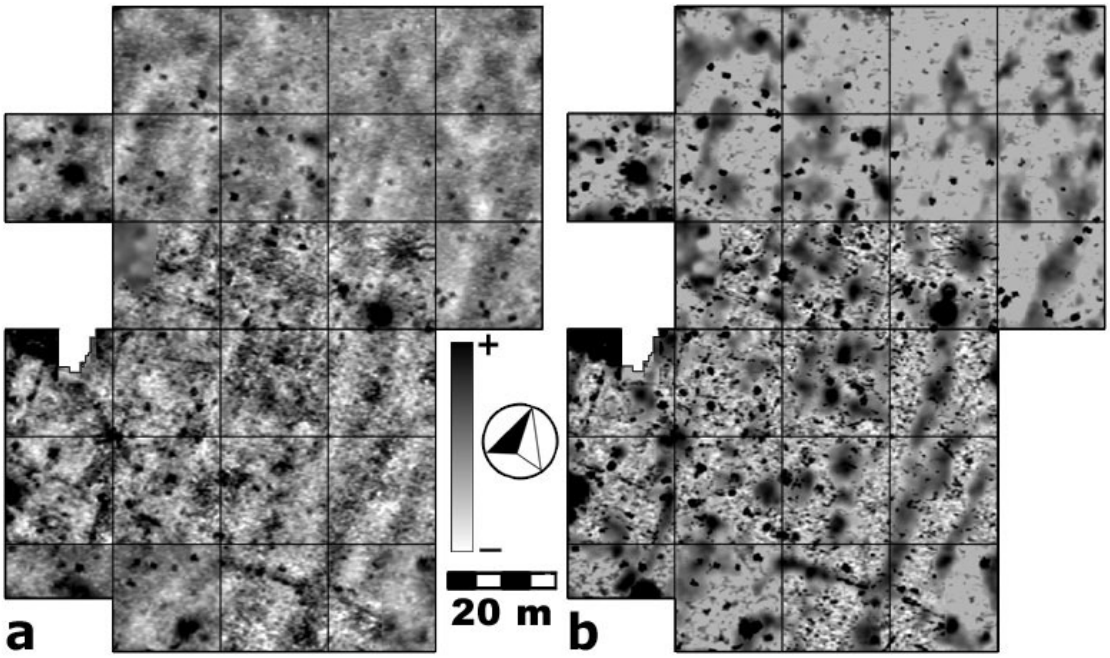


Figure 5.37. *Continuous fusion results at Silver Bluff Plantation. a) Product of GPR, magnetic gradiometry, and the inverse of electrical resistivity. b) Per-pixel maximum function of the same input layers.*

5.11.2.3. Cluster analysis

The unsupervised classification method of *k*-means cluster analysis (Section 5.5.7) was examined using the three geophysical layers from Silver Bluff Plantation. *K*-means cluster analysis identifies natural groupings of multivariate data by examining relationships between layers and isolating relatively homogenous regions. Any number of clusters can be found, but four was determined to be a maximum for these data, as additional clusters were either small or only sub-divided larger ones into ambiguous categories. In Figure 5.38a, Cluster 1 tends to show the background, or absence of anomalies; Clusters 2 and 4 describe increasingly more robust and negative resistivity anomalies; Cluster 3 primarily points to positive GPR anomalies. The spatial coincidence of Clusters 2-4, in effect, reveals that resistivity and GPR indicate similar subsurface phenomena, particularly in the central and southern reaches of the site. The rectangular grid-like distribution of anomalies is also emphasized. The magnetic gradiometry data (suggested previously to contain primarily noise) is very much underrepresented in this result, however, and was only visible when very high numbers of clusters were extracted.

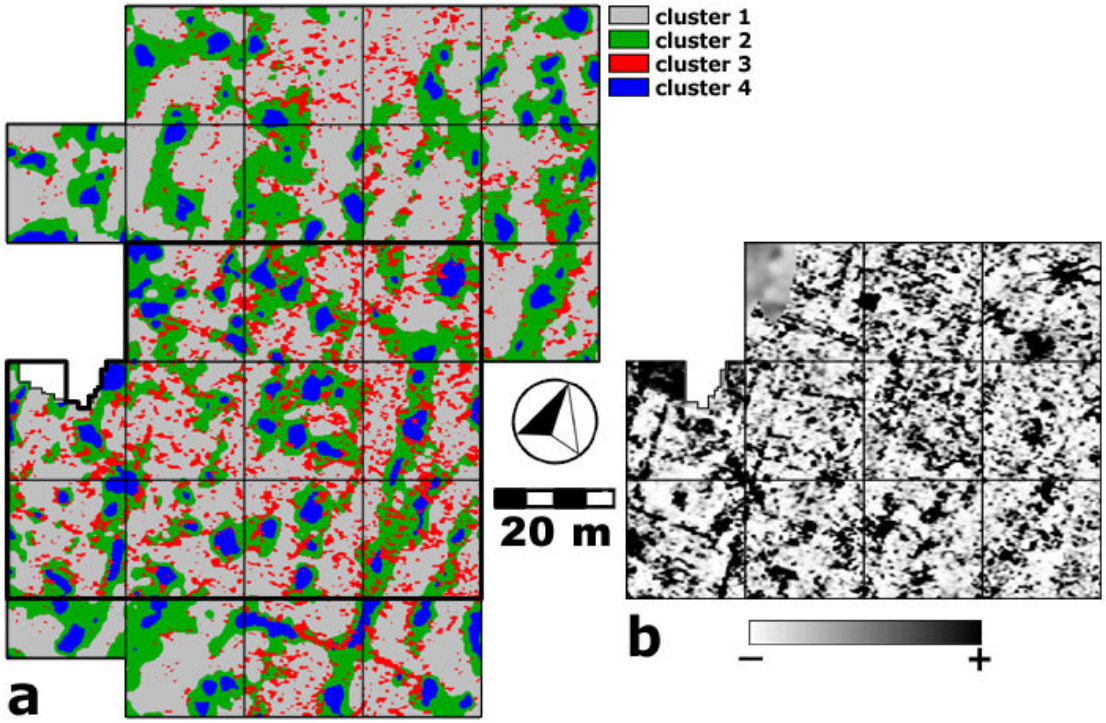


Figure 5.38. *Unsupervised and supervised classification results at Silver Bluff Plantation. a) K-means cluster analysis results for four clusters. b) Logistic regression probability surface for anomaly presence.*

5.11.2.4. Binary logistic regression

The supervised classification method of binary logistic regression (Section 5.7.3.1) was also explored using the Silver Bluff data. Only a subsection of the data covered by all three data sets was analyzed. The regression was calibrated using manually digitized training sites aimed at robust anomalies with geometric attributes suggestive of cultural features, for a total sample size of n=778 (out of N=307,200). The following regression equation was derived:

$$L = -2.136 - 0.552*RES + 1.926*GPR + 0.442*MAG$$

showing that the GPR data dominates the discrimination, followed by resistivity and finally magnetic gradiometry. A pseudo-R^2 = .29, indicating fairly good fit (Clark and Hosking 1986). Using the logistic transformation, $p = (1+exp(-L))^{-1}$, the regression function is converted to a 0-1 scale, interpretable as a probability surface for anomaly presence when mapped (Figure 5.38b). The mapping reflects the large weight of GPR because it looks very similar to the original GPR time-slice (see Section 4.6.3).

5.11.3. Kasita Town, Georgia

Geophysical data available from Kasita Town were limited to magnetic gradiometry, electrical resistivity, and ground-penetrating radar. In preparation for data integration each layer was resampled to a common resolution of .25 x .25 m and standardized to have a mean of zero and a standard deviation of one. In addition, the

upper five GPR depth slices were integrated by principal components analysis, with the first component explaining 43 percent of the variance (see Section 5.11.1.1). When compared against the individual slices, the first principle component (Figure 5.39) appears to show the majority of anomalies, so it was used for all subsequent integrations presented in this section. This insured that a large number of GPR time-slices would not dominate integration results. On Figure 5.39 and subsequent results, the strong linear anomaly that bisects the study area represents a recent pipeline. The GPR data offer enough detail that a number of truck tracks may also be seen in the northwestern part of the study area. These features were initially discussed in Section 3.3.4.4.

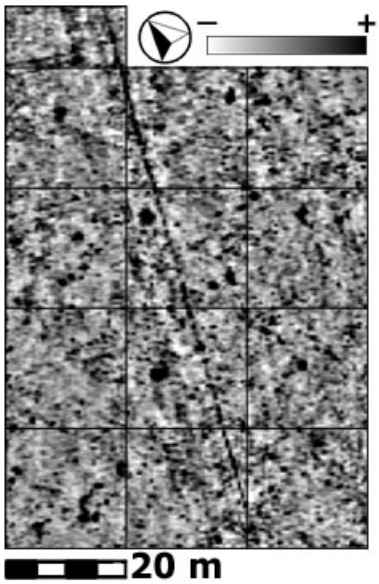


Figure 5.39. *Integration of GPR depth slices by principal components analysis at Kasita Town. Five GPR slices representing .2 m depth increments for the upper 1 m were used. This is the first principal component and represents 43% of the total variance.*

A correlation matrix between the three data layers shows them to be uncorrelated (Table 5.15), prohibiting use of such integrating methods as principal components analysis (Section 5.7.1). As with data fusions in other project sites, the magnetometry data were sometimes converted to absolute value before integration. Similarly, the electrical resistivity data were sometimes inverted because low resistivity anomalies corresponded more closely with positive ones in other data layers.

Table 5.15. *Correlation matrix between geophysical data layers at Kasita Town.*

	GPR PCA C1	Magnetic Gradiometry	Electrical Resistivity
GPR PCA C1	1.00	.0138	.0133
Magnetic Gradiometry	.0138	1.00	.0359
Electrical Resistivity	.0133	.0359	1.00

In other project sites there were many clear patterns in anomaly form and distribution identifiable in all data layers, as should be expected in archaeological sites containing cultural constructions (i.e., architectural and other built remains; see Section

4.6.4). Such patterned anomalies were used as indicators when experimenting with data integration methods, with results that enhanced their visibility generally favored over those that obscured them. In the Kasita Town data, however, there were few anomalies identifiable as cultural by form or distribution alone, so integration results were more difficult to judge. Reasons for this were unknown at the time of the geophysical surveys or during subsequent data analyses and fusions. They became clear in the excavation and validation phases of the project, however, where it was learned from numerous soil profiles that much of the site had been mechanically stripped and leveled in the recent historic period, with A-B soil horizons (and the archaeological deposits within them) missing or markedly truncated through much of the study area (see Section 4.6.4). The following material should be viewed with this in mind. The presented fusions were constructed under the assumption of largely intact cultural deposits at a major archaeological site of the early historic period, but it is small wonder that few apparent cultural patterns appear in the results.

5.11.3.1. Binary data integrations

Binary versions of each data set were created using arbitrary thresholds that appeared to isolate potential cultural anomalies, including low rather than high electrical resistivity anomalies and the absolute value of magnetic gradiometry. Since the data are statistically independent, it is no surprise that the addition of all three binary layers results in little overlap (Figure 5.40a). Fifty-eight percent of this result is background (i.e., not anomalous in any data set), with 35 percent in Class 1 (anomalous by only a single geophysical method) and seven percent in Class 2 (where at least two data sets simultaneously indicate anomalies). Less than one percent of the entire image contains anomalies indicated by all three layers (Class 3). Perhaps a more informative way of showing the binary data is through an RGB color composite (Figure 5.40b). The same information is revealed, but primary colors and their combinations reveal each anomaly's geophysical origin. Overall, these integrations of binary data are a useful way to clearly present the majority of anomalies. They are perhaps a useful first view of the integrated data because patterns difficult to see in the individual data layers (Section 5.5) are made more visible.

5.11.3.2. Continuous data integrations

Use of data in their original, continuous format can provide a richer view of results, with connections between robust anomalies preserved, but with more subtle results also shown. A constant of 100 was added to each data set (to insure that all values were positive), and the product of GPR, the absolute value of magnetic gradiometry, and the inverse of electrical resistivity was computed. This result effectively emphasizes most robust anomalies (Figure 5.41a). A distinctively different result is achieved by mapping magnetic gradiometry in red, covered by translucent green-tinted resistivity and blue-tinted GPR images (Figure 5.41b). While both of these integration methods are useful, in this case it appears that the latter result represents much more of the total variation in the data.

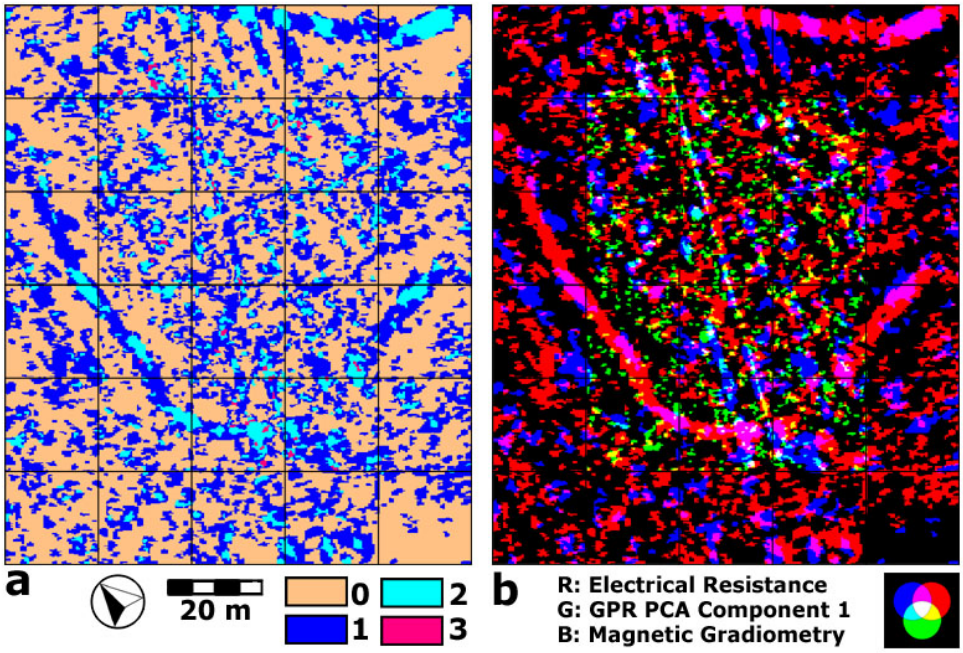


Figure 5.40. *Binary data integrations at Kasita Town. a) Sum of binary layers. b) RGB color composite of individual binary layers.*

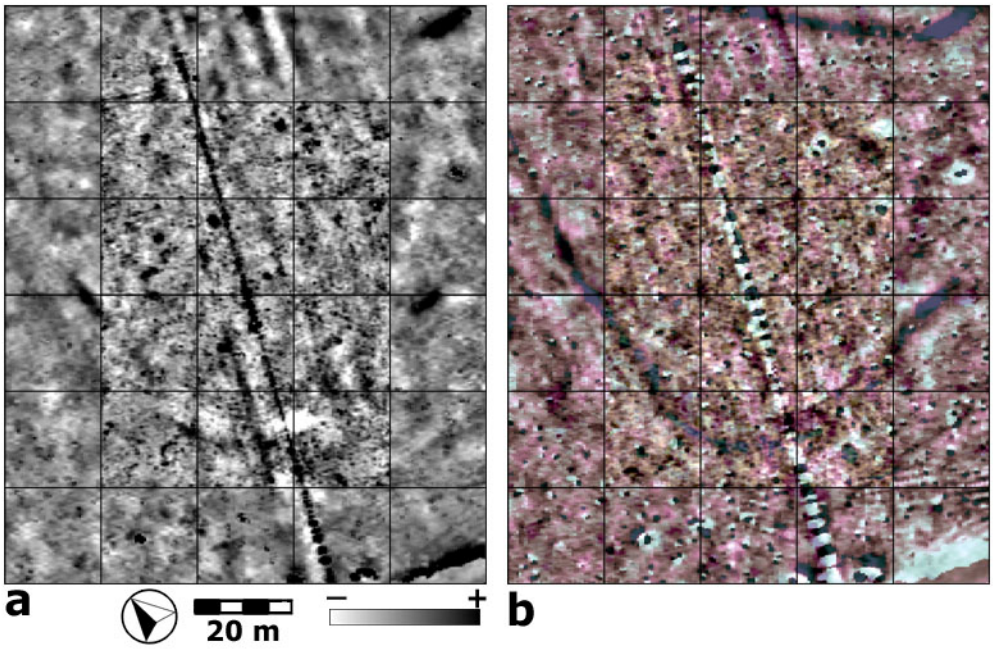


Figure 5.41. *Continuous data integrations at Kasita Town. a) Data product. b) The maximum of magnetic gradiometry (red), electrical resistivity (green), and GPR principle component 1 (blue).*

5.11.3.3. Binary logistic regression

Binary logistic regression was undertaken using the three continuous data sets in their standardized form (but absolute value was not taken). Training sites were selected

by a five percent random sample of those areas indicated by at least two binary geophysical data sets (Classes 1-2 in Figure 5.40a), resulting in 800 sample points. The analysis yielded pseudo-R^2=.16, and the following function:

$$L = -0.840 + 0.619*GPR + 0.249*MAG - 1.151*RES.$$

The coefficients indicate the relative contribution of each layer to the model, showing that electrical resistivity received almost twice the weight as GPR, and magnetic gradiometry less than a fourth as much weight. Figure 5.42a maps the above function. The logistic transformation, $p = (1+exp(-L))^{-1}$, was applied to the function, and the resultant 0-1 scale was then reclassified into five equal intervals (Figure 5.42b). The relatively low pseudo-R^2 indicates the distribution of the training sites is not well modeled by the geophysical layers, suggesting an ineffective fusion. Though this is difficult to judge qualitatively owing to a lack of apparent cultural anomalies that could be used as indicators, it appears that the final result does not include some of the subtle magnetic anomalies that are easier to visualize in both discrete fusions (Figure 5.40) and the translucent color overlay (Figure 5.41a).

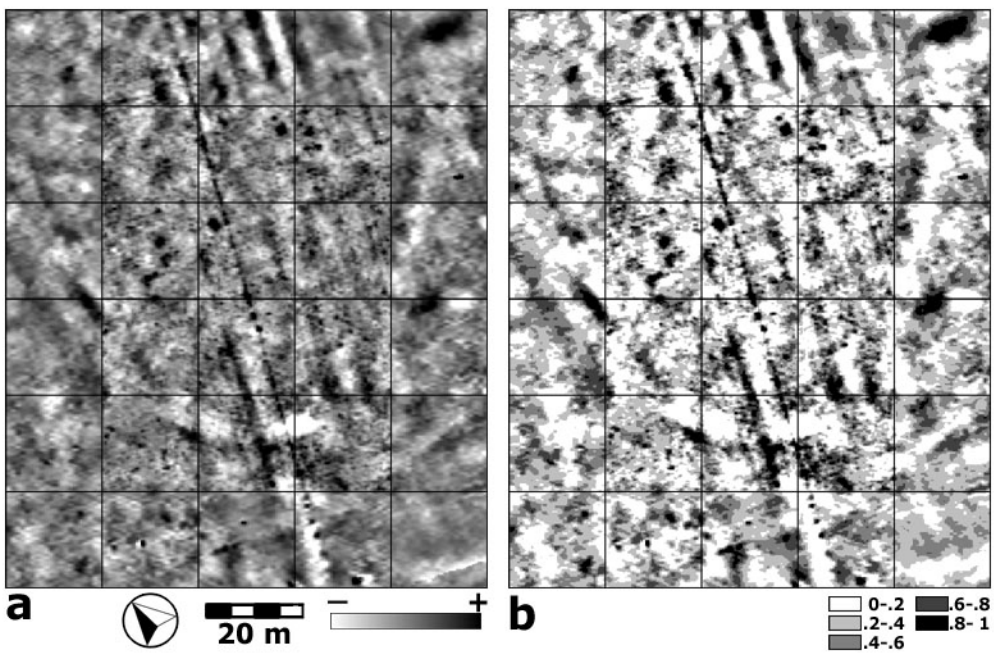


Figure 5.42. *Binary logistic regression results at Kasita Town. a) The raw regression function. b) The regression function after a logistic transformation and classified into five equal-interval categories.*

5.11.3.4. K-means cluster analysis

Perhaps a better method of fusing results from Kasita Town lies in unsupervised classification, in hope that statistically-defined natural groupings are more meaningful than subjective interpretations of data lacking any clear cultural patterns. Several iterations using different numbers of clusters were examined using the unaltered geophysical data layers. K=5 clusters was found to be an effective maximum, since any

additional clusters were small in area and difficult to interpret. The result (Figure 5.43) appears to isolate the major magnetic gradiometry and electrical resistivity anomalies, with less emphasis on those indicated by GPR. Cluster 1 includes areas of low electrical resistivity, positive GPR anomalies, and average magnetic gradiometry. Cluster 2 only seems to show high and low magnetic gradiometry anomalies, but not subtle ones. High electrical resistivity anomalies make up the remaining clusters, in which Clusters 5, 3, and 4, respectfully, make up successively more robust anomalies.

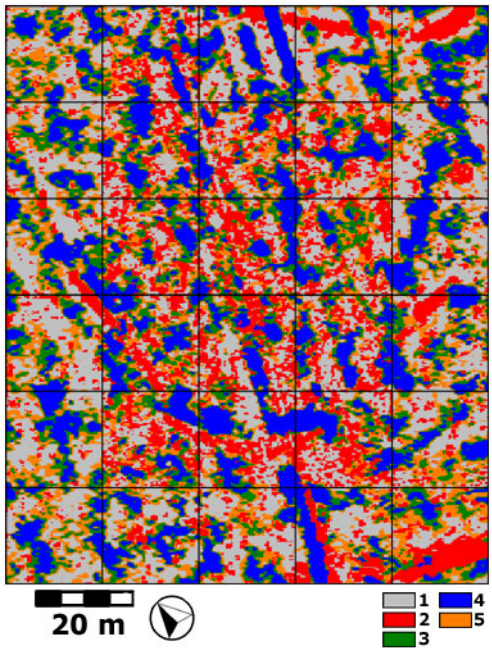


Figure 5.43. *K-means cluster analysis results for five clusters at Kasita Town.*

5.11.3.5. Conclusions

The Kasita Town geophysical data are difficult to interpret because they lack recognizable cultural patterns as a basis for visual evaluation of fusion-integration results. Findings are therefore judged by the ability of a fusion to preserve as much of the input data layers as possible. The RGB composite of binary layers (Figure 5.40b) seems to most effectively accomplish this task, displaying all three layers simultaneously while not excluding a significant portion of any one data set. The sum of binary layers provides a very similar result (Figure 5.40a), but with less information because the ability to identify an anomaly's source is lost. The use of translucent overlays of the original continuous data sets (Figure 5.41b) is also an effective means of displaying the majority of the information simultaneously, with the added benefit of preserving subtle anomalies that are at least partially lost in binary versions. Supervised and unsupervised classifications were not as effective at Kasita Town as they were at other sites. Binary logistic regression (Figure 5.42) produced a result with a very low pseudo-R^2 that did not well illustrate magnetic gradiometry information. *K*-means cluster analysis effectively identified high and low electrical resistivity and robust magnetic gradiometry anomalies, but did not isolate subtle magnetic gradiometry anomalies and appears to have lumped positive GPR anomalies into the background (Figure 5.43). These results together seem to suggest that poor data (lacking geometric patterns characteristic of cultural activity) yields poor

statistical-based fusions. In such cases, however, image-based fusions such as RGB composites and translucent overlays may be very useful ways of effectively displaying the bulk of information from three input layers.

5.12. INTERPRETATIONS AND VECTORIZATION

Kenneth L. Kvamme and Eileen G. Ernenwein, University of Arkansas

5.12.1. Overview of Process

The production of a final "end-image" that clearly portrays subsurface features of possible cultural origin through extensive data processing and enhancement (e.g., Section 4.4) is not typically the ultimate goal of remote sensing. The utility of results is greatly benefited by an analytical phase that transcribes interpreted features accurately onto maps. Transforming information in remotely sensed imagery into a record of the archaeology is complicated, but important. It requires knowledge of the kinds of archaeological features—from all possible cultures and times—that might occur in a site, experience in their remote sensing signatures, which can vary regionally according to soils, geology, and climate, knowledge of the vast body of theory that underlies each remote sensing domain, the amount of regular or geometric patterns expressed by anomalies, and technical knowledge of cartography and mapping. Many of these topics were reviewed in Section 4.5.

Transcribed maps are composed of points, lines, and polygons—*vectors* in GIS parlance—that represent significant anomalies interpreted in the processed raster imagery of the various remote sensing products. Interpretation can therefore be viewed as a raster-to-vector conversion process, because the continuous changes witnessed on a pixel-by-pixel basis in geophysical results must be classified to specific categories of archaeological types and made discrete. Interpretations are built up, and cumulate, through inspection of *each* remote sensing data source for a given piece of ground. Some features may be revealed only by a single sensor, while others are visible in several data sets, allowing their presence to be verified, expanded, and ultimately codified and classified. GIS is particularly useful in this process because recognized archaeological features may be digitized on-screen directly into a map coordinate system and image manipulation tools can locally enhance or highlight features to facilitate their recognition. With transcribed vectors entered into regional GIS-driven databases, a permanent inventory of known or suspected archaeological features is developed, essential to guide management and planning decisions and future work (Bewley 2000:8; Wilson 2000:225-235).

Several of the foregoing requirements posed difficulties to the project. In particular, success in the interpretive process requires considerable knowledge of the nature of the archaeology anticipated in a site and corresponding geophysical responses in order for cultural anomalies to be identified. Although the project team had considerable experience in archaeological geophysics from throughout the country and was well versed in the method, theory, and instrumentation, years of experience in a region are probably necessary to become sensitive to the nature and range of archaeological occurrences and the character of the anomalies they produce. As a hedge, local expert consultants in each region were hired and extensive reviews of regional literature and prior archaeological work was made.

5.12.2. Methods

As indicated in Section 4.1.2, anomalies are of two general types—those stemming from cultural processes and those arising from natural circumstances. The former includes the targets of potential interest to the archaeologist. The task, then, is to weed out the cultural from the natural, and within the former, define those that might be significant to archaeological investigations. In other words, sites might include many anomalies representing cultural features or deposits, but only some may be relevant to the archaeological occupation or component being examined. For example, many metallic anomalies are of cultural origin, but most may represent modern rubbish. Plow marks and historic field boundaries are also a cultural phenomenon, but usually represent relatively recent impacts to a site, and so are not of interest. The same may be said of recent landscaping practices, pipelines, roads, trails, and other elements that commonly cause anomalies in geophysical data sets.

Generating geophysical interpretations is a multi-step process that makes use of the general techniques outlined in Section 4.5. First, background research, local experts, and experience in a region identify the kinds of archaeological features that can be expected in the site. Second, prior work, geophysical theory, and experience reduces this initial set to a list of archaeological features that *might* be visible in the data given the geophysical techniques and sampling densities employed and environmental conditions. This involves the sizes or volumes of the potential archaeological features, their physical makeup, and the amount of regular pattern they might exhibit. Third, with this background the geophysical data are examined to associate specific classes of observed anomalies with as specific as possible archaeological types. For example, at the most general an anomaly might simply be classified as a "linear cultural feature"; at a more specific level it might be determined to be a "brick wall" (depending on the nature and detail of the geophysical evidence and prior knowledge of the site). Of course, there always remains a collection of amorphous anomalies that can only be identified as *possible* archaeological features, due to their ambiguous signatures in the data. Finally, the last step defines likely archaeological features by encoding them as discrete point, line, or polygon entities via GIS.

This last task is facilitated by GIS software. Color palettes and contrast can be changed and manipulated to exaggerate anomalies of interest or present alternative renderings. Reclassification tools can be employed at various thresholds to define anomalies at various levels of robustness. Other geophysical data sets can be "blinked" on and off to rapidly investigate additional evidence or corroborate findings. Data fusions can merge several lines of evidence to make weak anomalies more robust or to show anomalies in a larger context. A number of pitfalls are associated with vectorizing geophysical data. The continuous nature of the data often means that anomaly boundaries may not be as sharp and discrete as we would like—boundaries can change as the contrast or palettes are manipulated. This makes it difficult to approximate the "true" boundaries of features and can cause inconsistently drawn with changes in settings. The vectorizing process is illustrated using Army City data below.

5.12.3. Evolving Perspectives

The initial site subjected to interpretation and validation through field excavations was Army City. That site, dating to 1917-1922, contains many elements of our own culture, and several historic photographs are available (Section 3.3.4.1). This made interpretation relatively "easy" compared to the other prehistoric or proto-historic sites of the project. Because we felt we knew or could recognize the sources of many anomalies at Army City, cultural feature types were defined in many classes and a relatively high amount of detail. While the field-testing generally proved successful (Section 5.16.1), it became obvious during fieldwork that the classifications were over-specified and too detailed, raising implications for accuracy assessments in the validation phase. If one is interpreting broad area anomalies as "floors" because they are near the main street of Army City where we know by photographs that shops were once present, and excavation shows a broad lens of gravel, is the interpretation correct? In one sense, we know the gravel caused the anomaly, but does gravel constitute a floor? Perhaps that gravel was once a substrate below a floor of tiles; on the other hand, perhaps that gravel was derived from the rubble of a fallen wall when the town burned and was later dismantled? With 20-20 hindsight, it is now believed anomalies should be placed in the most general of classes unless one has a great deal of experience in a region or at a particular site. Given the formation processes of Army City, from its construction and short life history through its demise by fire and dismantling (Section 3.3.4.1), such area features should have been more properly classified merely as "broad area anomaly of likely cultural origin" (with footnotes suggesting the possibility of floors or rubble lenses from the town's destruction). This over-specificity of course impacts accuracy rates when comparing assigned anomaly types to actual archaeological circumstances in the testing and excavation phase (Section 5.16).

Owing to the project's fast-paced timetable, there was no time to digest knowledge gained from the experience of the fieldwork at Army City in mid-November 2004, with the required SERDP annual meeting in early December, end-of-year Holidays, and the preparation of an anomaly classification and sampling design for fieldwork at Pueblo Escondido in late January 2005. Consequently, and unfortunately, the anomaly classification system at that site could not benefit from these lessons, with over-specification of anomaly types made once again with the same issues regarding classification accuracy. By the time field-testing was undertaken at Silver Bluff and Kasita Town in March and April of 2005, these lessons were appreciated and a revised and more generalized system of classification was in place, with fewer specific anomaly types defined. The following sections reflect this change in orientation from specific to more generalized anomaly classes.

5.12.4. Army City

Given the large number of remote sensing modalities acquired at Army City, with six different sensors imaging the site (Section 4.6.1), plus the extraordinary number of data fusions (Sections 5.3-5.10), a tremendous amount of information is available to define significant anomalies at Army City and classify them into meaningful cultural types. In addition to this information, a variety of historic data sources are also available. These include an historic plat map that defines the street and alley system of Army City

and several historic photographs. The plat map is important because it shows the spacing of streets and indicates the presence of alleys within every block, a circumstance verified in the photography (Figure 3.2). Knowledge of the existence of alleys helps explain a number of prominent linear anomalies. Although the plat map illustrates an ambitiously planned layout, development of Army City during its short life was generally confined north of the railroad (Figure 3.1; except for a small district for "Negro troops" south of the tracks), with few buildings outside of the commercial core (Figure 3.2; Hargrave et al. 2002). Historic photography usefully shows sidewalks, street gutters, open parks, gardens, and numerous structures, allowing correlation of numerous anomalies with these features. Moreover, the photography allows identification of the sources of very specific anomalies. One photograph reveals a large external chimney outside a wooden structure, explaining a robust anomaly reflecting its base (Figure 3.2, right background). Other photos show the burning Orpheum Theater and burned remnants of the Hippodrome, explaining robust magnetic anomalies in those areas.

The process of identifying anomalies that represent significant cultural features was accomplished through careful examination of the historic photographs and maps, each remote sensing modality, and the many data fusions. In the geophysical data, great reliance was placed in pattern recognition concepts, where anomalies of anthropogenic origin generally exhibit regular geometric shapes (Section 4.5.3). In other words, straight-line features, rectangular and square areas were interpreted as culturally significant. By reference to relative positions within the site, their size, and use of the maps and photography, it could be determined whether a linear anomaly represented a street gutter segment or part of a building's wall. Likewise, rectangular or square anomalies could be interpreted as parts of rooms or floors at one scale, or the boundaries of city blocks at another.

As a start-point, and considering only spatial form, anomalies were placed into two fundamental types: polygons (those covering areas) and linear features (the line data type). "Point" anomalies were encoded as very small polygons. Polygons represent interpreted features that cover areas, such as floors of structures. Lines represent the many lineations throughout the site interpreted as walls and street gutters. A second basic distinction was between robust and subtle anomalies. The former were defined as those with large measurements by a single sensor (usually beyond two standard deviations of the mean) or those simultaneously revealed by multiple sensors. Subtle anomalies could be visually distinguished from the background, but did not meet definitions of robustness. Anomalies were initially digitized singly for each sensor (Figure 5.44a,b), overlaid and consolidated (Figure 5.44c), compared against various data fusions (Figure 5.44d), and reconsolidated again in an iterative process with aims to merge, generalize, and simplify. After many iterations and revisits to the data, sixteen classes of anomalies were ultimately defined, plus a background class for a total of 17 classes. These anomaly categories are illustrated in (Figure 5.44e) and are described in Table 5.16.

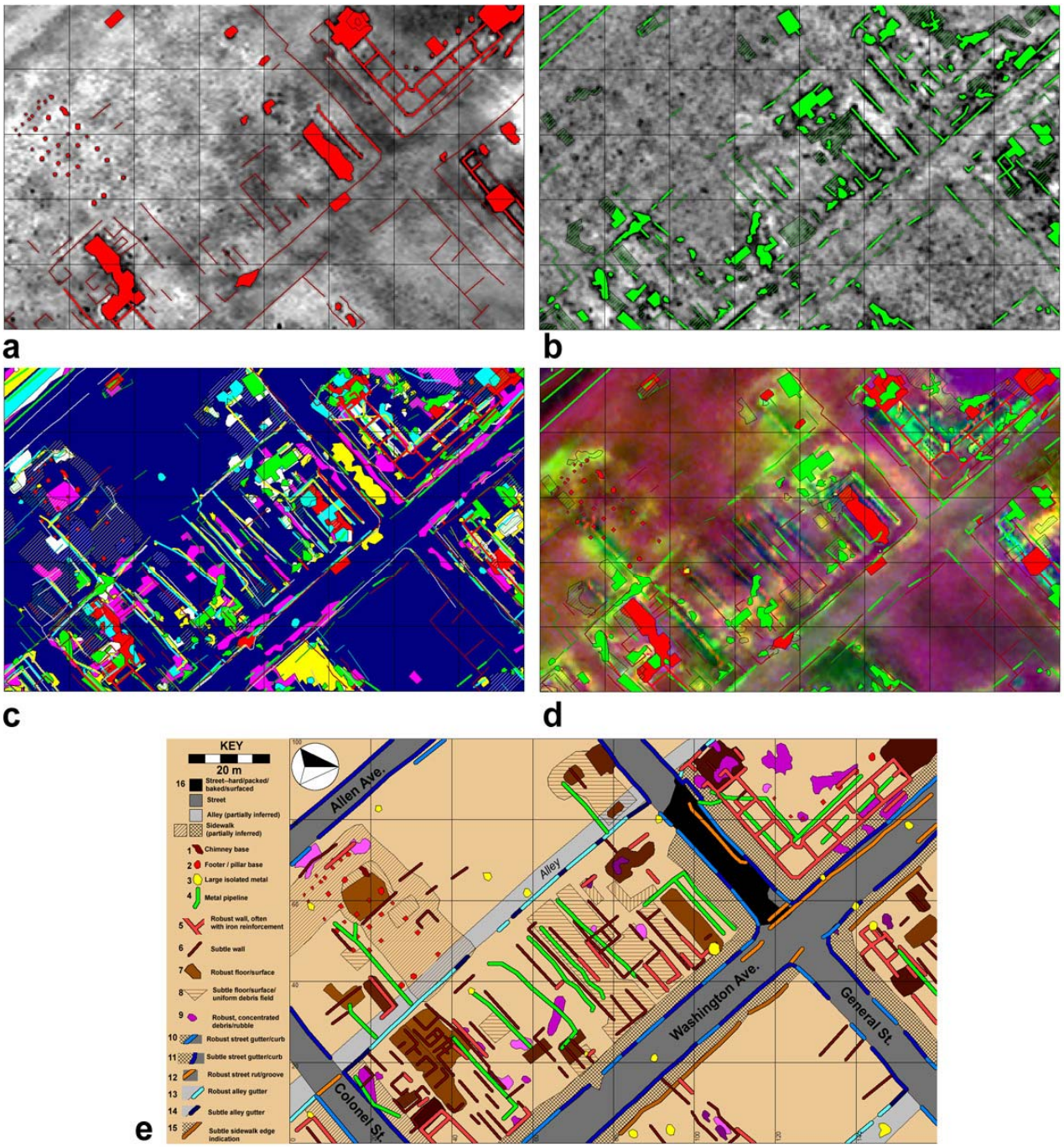


Figure 5.44. *Illustration of the process of vectorizing anomalies at Army City. a) Initial interpreted resistivity vectors, classified as robust and subtle by color brightness differences. b) Initial interpreted GPR vectors, classified as robust and subtle by color brightness differences. c) Initial interpreted vectors for all sensors overlaid. d) Resistivity and GPR vectors overlaid on fused imagery that also portrays magnetic susceptibility and conductivity anomalies. e) Final simplified and consolidated vectors representing 16 classes of anomalies.*

As can be seen in Table 5.16, a rather eclectic mix of classified anomaly types is offered.

Table 5.16. *Description of 17 defined anomaly classes at Army City.*

Class	Name	Line / Polygon	Description (*and principal source of indication*)
1	*Chimney base*	P	A robust irregular anomaly that an historic photo suggests was the locus of a chimney (*Res*)
2	*Footing*	P	Building footing / support or possible pillar base. Robust, regularly distributed anomalies of small size (< $1m^2$) (*Res*)
3	*Isolated metal*	P	Robust anomaly of small point size or dipolar form suggesting isolated iron artifact of large size. Located away from buildings where metallic anomalies might point to wall construction or pipe metals. (Mag, MS)
4	*Pipe*	L	Robust linear anomalies of dipolar form (*Cond, Mag, MS*)
5	*Debris*	P	Robust, distinct, irregularly shaped anomaly (*GPR, Mag, MS, Res*)
6	*Robust wall*	L	Linear feature at least 1 m thick, probably concrete, and generally iron reinforced (*GPR, Mag, MS, Res*)
7	*Subtle wall*	L	Narrow linear feature of moderate magnitude, often vaguely indicated
8	*Robust floor*	P	Broad area/surface of elevated measurements with indication of linear boundary and rectangular shape. Packed earth, concrete, brick are possible (*All methods*)
9	*Subtle floor*	P	Broad area/surface of elevated measurements with indication of linear boundary and rectangular shape. Packed earth, debris field, burned surface (*All methods*)
10	*Street surface*	P	Broad segment of General Street with distinctly elevated measurements suggesting greater compaction, a possible surface, or intensive firing (*MS, Res*)
11	*Robust street gutter*	L	Robust linear feature at boundary of loci of principle streets. Historic photos indicate distinct drainage gutters with pooled water as well as an offset concrete sidewalk edge. Indications of one or both may be present (*All methods*)
12	*Subtle street gutter*	L	Subtle linear feature at boundary of loci of principle streets. Historic photos indicate distinct drainage gutters with pooled water as well as an offset concrete sidewalk edge. Indications of one or both may be present (*All methods*)
13	*Street rut*	L	Robust linear features within Washington Street and elsewhere. Historic photos indicate large ruts in Washington Street (*GPR, Mag, MS*)
14	*Robust alley gutter*	L	Pronounced linear features at the loci of alleys suggest a gutter (*GPR, Mag, MS*)
15	*Subtle alley gutter*	L	Subtle linear features at the loci of the alleys suggest a gutter (*GPR, Mag, MS*)
16	*Subtle sidewalk edge*	L	Subtle linear features inset from street edge suggest locus of sidewalk edge (*GPR, MS*)
17	*Background*	P	Locations without anomalies

Abbreviations: *Cond* (Conductivity), *GPR* (Ground-penetrating radar), *Mag* (Magnetometry), *MS* (Magnetic susceptibility), *Res* (Resistance), *Therm* (Thermal)

They can be grouped into several categories. These categories include “point” anomalies (polygons of small size) in (1) chimney base, (2) footing, and (3) isolated metal categories. Metal artifacts form a second category that includes (3) isolated metal and (4) pipes. Structural anomalies that cover areas represent a third grouping that consists of bounding linear features of (5,6) robust and subtle walls, and area features most likely representing (8,9) robust and subtle floors. The fourth and largest category of anomalies are associated with streets and include (10) prepared street surfaces, (11,12) robust and subtle street gutters, (13) street ruts, (14,15) robust and subtle alley gutters, and (16) subtle sidewalk edges. The final class (17) represents the background lacking anomalies. These are the classes ultimately tested by excavation in the anomaly validation phase, described in Section 5.16.1.

5.12.5. Pueblo Escondido

Vector interpretation of Pueblo Escondido was facilitated by the high quality GPR, magnetic susceptibility, and magnetic gradiometry data. Prior knowledge of the site in the form of site reports, research articles, and publications describing comparable sites in the area further aided in understanding the architectural configurations that could be expected. Clearly rectangular and linear anomalies in many of the geophysical data sets made interpretations more reliable than other sites where only small irregularly shaped anomalies are present. Initial interpretations of the geophysical data were discussed in Section 4.6.2. Pueblo Escondido has been documented by the El Paso Archaeological Society (EPAS) (1968), the University of Texas at Austin (1975), and Fort Bliss (1975). Findings from the El Paso Archaeological Society were published in a brief article (Hedrick 1967). Together, these reports describe Pueblo Escondido as the largest El Paso Phase site in the area, and having at least four rows of pueblo room blocks as well as several extensive trash middens, lag deposits of artifacts and fire-cracked rock, and fire hearths (EPAS site form 1968). This information combined with publications describing other similar sites provided a baseline for cultural interpretation of the geophysical data.

Interpretations focused on the large contiguous block of geophysical data designated as Area C (see Figure 3.4), because these data were collected using optimal sampling densities and instrument configurations that were determined after an initial exploratory phase of data collection (Areas A and B). Area D (Figure 3.4) was not yet surveyed with geophysical instruments at the time of this initial vector interpretation. Of the four data sets collected for Area C, only GPR, magnetic gradiometry, and magnetic susceptibility are used. The fourth data set, conductivity, appears to show broad variations in soil properties with little or not relationship to cultural features (see Section 4.8.7).

Previous experiences with interpretation and subsequent excavations at Army City prompted the SERDP team to utilize less specific anomaly classes, and a somewhat more generalized vector interpretation was created for Pueblo Escondido. A preview of the geophysical data (Section 4.6.2) shows several rectangular and linear anomalies that are likely to be cultural in origin, as well as countless amorphous anomalies.

The architecture of El Paso Phase villages typically includes linear alignments of rectangular pueblo room blocks oriented east-west, each one containing six to ten medium-sized rooms and one very large communal room. Individual rooms typically range in size from 8 to 20 square meters, while large communal rooms can be upwards of

50 square meters (O'Laughlin 2001). Smaller rooms usually have two primary roof support posts and a hearth near the southern wall, while larger rooms can have a support post in each corner and two or more hearths and other floor pits and features. This information, combined with the general knowledge of other typical archaeological features such as middens, hearths, and storage pits, was used to interpret the size and shape of anomalies in the geophysical data. Anomalies were further differentiated based on their geophysical signatures. For example, a large number of what appear to be individual rooms were revealed in the GPR data, but only some were indicated with magnetic gradiometry. Since the latter responds to remanent magnetic fields created during burning, it can be hypothesized that those anomalies having strong magnetic signatures were burned, and those with little or no magnetic signatures were not. It could also be ruled out that any lack of magnetic anomalies was due to depth, since the GPR data shows all anomalies within the upper meter of deposit. The magnetic gradiometry data were also used to differentiate between potentially burned and other small amorphous anomalies.

Anomalies were also classified based on associations with other anomalies and regularities in their spacing. For example, small amorphous anomalies could represent a wide variety of archaeological or non-archaeological features, but their locations in relation to probable house anomalies allow their assignment as "interior" versus "exterior" forms. Many exterior and small amorphous anomalies seemed to be randomly distributed about the entire survey area, while others were arranged in linear alignments paralleling the predicted rooms.

The data sets showed great variety in anomaly strengths. Robust anomalies often can be more confidently classified owing to solid and unambiguous forms. Some anomalies, however, are quite weak but still take a shape that is suggestive of cultural activity. It was therefore decided that each class of anomalies would be divided into "subtle" and "robust."

Based on these criteria, fourteen different classes of anomalies were defined and digitized based on shape, size, geophysical properties, robusticity, location, and regularity of form or spacing. A fifteenth class, labeled "background," was also created to represent all those areas where no features are expected based on the geophysical data. Table 5.17 lists these anomaly classes along with a brief description. The locations of all defined types are shown in Figure 5.45. This vectorized interpretation portrays an easily understood map of potential archaeological features. In general, it illustrates a typical but very large El Paso Phase pueblo village consisting of at least four and possibly nine discrete groups of rooms forming room blocks, as well as several isolated rooms that could represent partially eroded room block groupings or perhaps pithouses of a different architectural style. In addition, there are countless interior and exterior room features that could represent storage pits, hearths, roasting pits, middens, or perhaps burials.

Table 5.17. *Description of 15 defined anomaly classes at Pueblo Escondido.* Abbreviations: *GPR* (Ground-penetrating radar), *Mag* (Magnetometry), *MS* (Magnetic susceptibility)

Class	Name	Line/ Polygon	Description
1	*Room block floor*	P	Strong aerial anomaly within rectangular room block– significant surface by GPR or MS
2	*Magnetic room block floor*	P	Strong magnetic anomaly within rectangular room block–by Mag/MS – possible burning or igneous rock.
3	*Inferred room block floor*	P	No direct anomaly indication. Floor inferred in room block between walls.
4	*Outlying floor*	P	Medium-large (>2m) anomaly robustly indicated. Possible patio, storage room, or pithouse floor.
5	*Magnetic room block wall*	L	Linear anomaly surrounding room or room block, high magnetically (Mag or MS): burned or perhaps igneous foundation.
6	*Robust room block wall*	L	Strongly indicated linear anomaly surrounding room or room block, by GPR or MS
7	*Subtle room block wall*	L	Subtly indicated linear anomaly surrounding room or room block, by GPR, MS, or Mag
8	*Interior room block feature*	P	Circular or irregular-shaped feature in room by GPR (not magnetic area). Probably wall fall/rubble, possibly pit or post mold.
9	*Circular magnetic interior room block feature*	P	Circular anomaly 1-2 m in diameter within roomblock, by Mag. Likely hearth or high fired area by room burning
10	*Robust lineation*	L	Linear feature paralleling room blocks. Walls, boundaries, edge of packed use area? MS & GPR
11	*Subtle lineation*	L	Similar to above but subtle—deeper or smaller feature.
12	*Small polygon systematically distributed*	P	Small circular to irregular anomaly by MS or GPR, 1-2 m in diameter distributed in repetitive linear pattern, often paralleling room blocks. Storage pits, waste pits, posts, or work areas.
13	*Small polygon randomly distributed*	P	Similar to above, but randomly distributed.
14	*Circular magnetic exterior anomaly*	P	Circular anomaly 1-2 m in diameter exterior to room blocks. Possible hearth, roasting pit, or concentrated burn area.
15	*Background*	P	Locations without anomalies

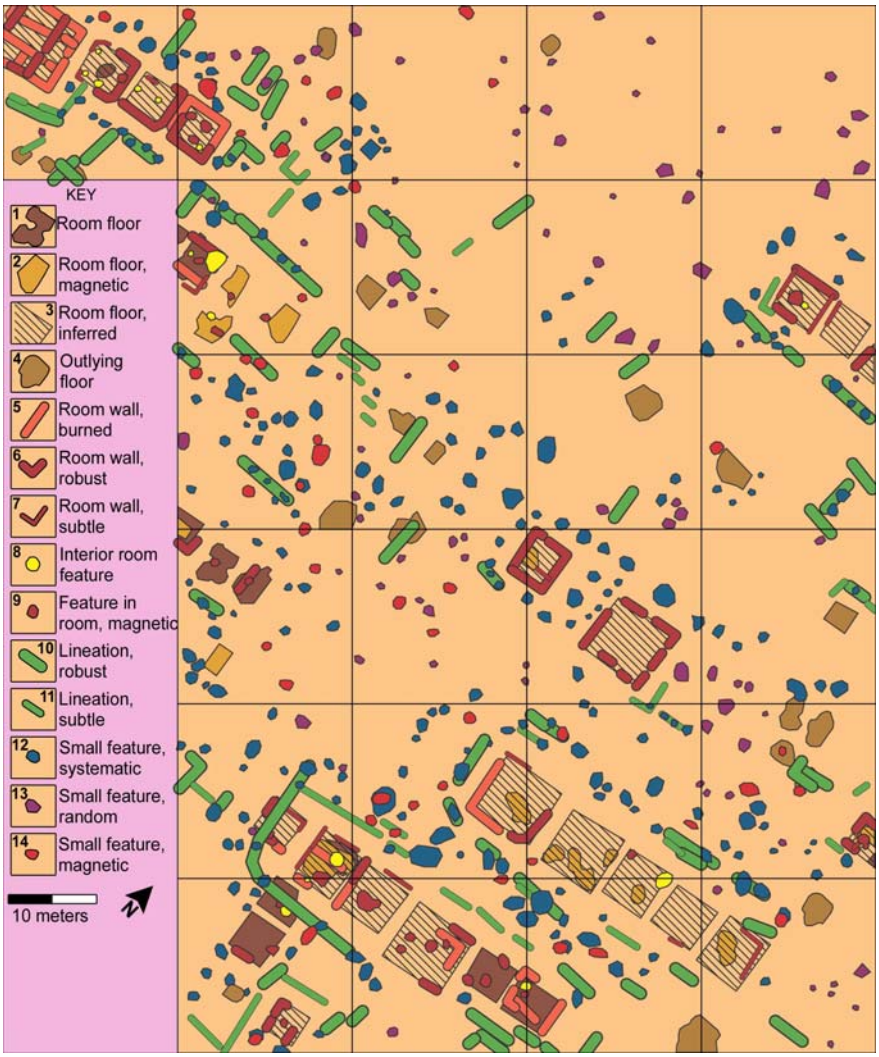


Figure 5.45. *Vector interpretation for Pueblo Escondido. Linear Anomalies are exaggerated in size to aid visibility.*

5.12.6. Silver Bluff

The South Carolina Institute for Archaeology and Anthropology (SCIAA) conducted small-scale excavations at this site through 2002, with the remains of nine buildings and a substantial stockade thus far documented (see Section 4.6.3). Extensive excavations by Herron and Moon (2005) indicate much about the character, depth and orientation of archaeological features, with substantial brick buildings to ephemeral slave quarters. Historically, the property began as a trading post that later evolved into a component of a major plantation with original buildings used for storage. Numerous cultural anomalies arising from constructions over the course of more than a century of intensive use may therefore be expected and are strongly suggested by the geophysical findings (Section 4.6.3). Unfortunately, the many pin flags used for prior archaeological investigations caused prominent anomalies that made most of the magnetic gradiometry data of little use (Figure 4.45a). Electrical resistivity survey revealed numerous positive and negative linear anomalies aligned on north-south and east-west axes (the same as excavated archaeological features) strongly suggestive of constructions (Figure 4.45b). GPR survey supported this result but with much greater specificity by indicating linear

and rectilinear anomalies with great clarity —some apparently defining individual rooms (Figure 4.45c-e).

In keeping with the evolving project perspective of less specific classifications of anomalies to archaeological classes unless there was near-certainty, anomalies were classified into a relatively small list of spatial-geophysical types against which several plausible archaeological causes could be listed (Table 5.18). For example, the GPR data reveal linear and rectilinear anomalies of clear cultural origin in several slices (Figure 4.45c, d). While most could have been labeled as "walls," they are listed only as robust or subtle GPR anomalies, with the understanding they could represent an actual wall (e.g., of brick), a line of closely-spaced palisade posts, a builder's trench that once held posts, or a series of post holes at regular intervals. On the other hand, a number of classes were rather specific, including "historic iron or steel" which likely points to pin flag remnants, but could conceivable indicate the loci of historic iron artifacts. The "lightening strike" class (Table 5.18) was pointed out by ArchaeoPhysics LLC, the geophysical contractors who acquired the data. Located under a prominent large tree evidently struck by lightening, a GPR anomaly of characteristic form (Figure 4.45c, d) shows RDP changes caused by the massive voltages and temperatures (see Jones and Maki 2005).

Table 5.18. *Thirteen anomaly types at Silver Bluff Plantation. All anomaly classes treated as polygon data.*

Class	Name	Description
1	*Historic iron or steel*	Historic iron or recent steel (primarily pin flag) artifacts
2	*Subtle linear magnetic*	Possible a wall of magnetic or burned material, or an incised trail filled with magnetically enriched surface soil
3	*Magnetic point in linear alignment*	A series of closely spaced magnetic anomalies that appear in a row and could indicate fired soils surrounding burned posts
4	*Linear negative resistance*	Roads, paths, or lanes, possibly "sunken"
5	*Linear positive resistance*	Possibly mounded "berms" adjacent to principal roads, paths, or lanes, or bounding the environs of a structure
6	*Area negative resistance*	Open space between lanes or structures?
7	*Area positive resistance*	Floors? Compacted spaces?
8	*Subtle linear GPR*	Walls, builder's trenches, closely-spaced post holes
9	*Subtle negative linear GPR*	Space between walls/structures? Lane between walls/structures?
10	*Robust linear GPR*	Walls, builder's trenches, closely-spaced post holes
11	*Area GPR*	Floors? Compacted spaces?
12	*Area robust GPR--deep*	Larger road, drainage, paleochannel?
13	*Lightening strike*	Lightening strike

These anomaly types are mapped in Figure 5.46. The GPR and positive resistivity anomalies emphasize the many lineations and rectangular organization characteristic of much of the site. The several linear magnetic anomalies parallel this distribution. Spaces between these lineations tend to be illustrated as negative resistivity anomalies, perhaps showing the lanes or open spaces between developed areas (compare Figure 4.45).

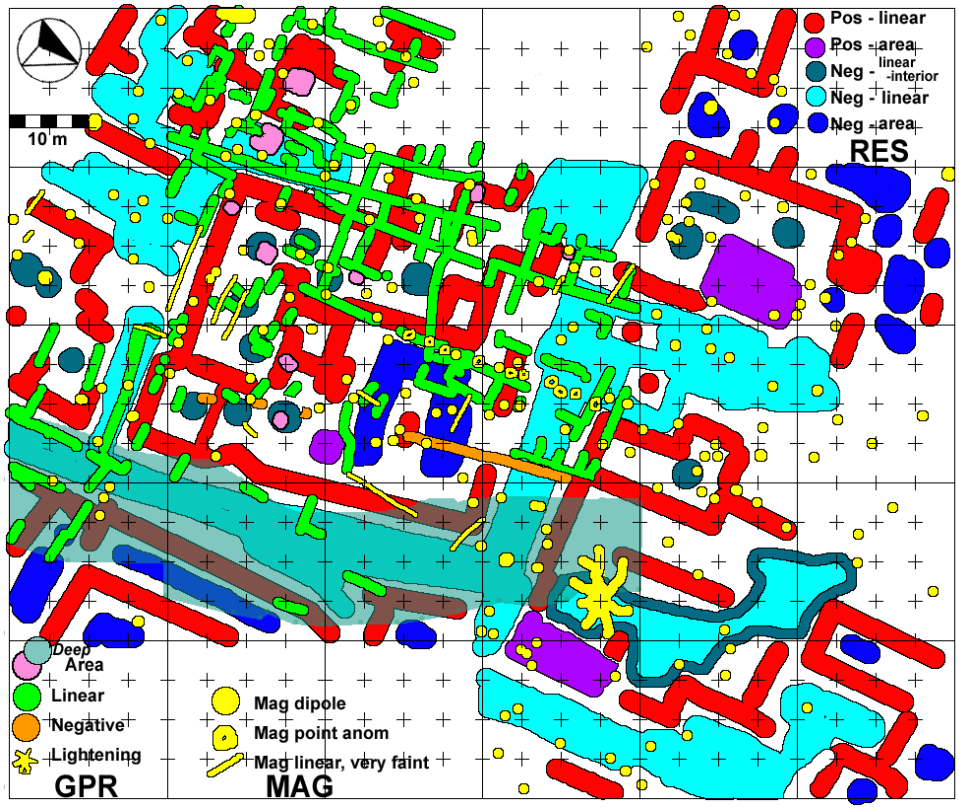


Figure 5.46. *Mapping of the classified anomaly types for Silver Bluff Plantation. Widths of linear GPR anomalies are greatly exaggerated.*

5.12.7. Kasita Town

Previous excavations at Kasita Town and similar sites provided a baseline for interpretation of geophysical data (see Section 3.3.4.4). Excavations in preparation for the construction of Lawson Airfield by Gordon Willey in 1936 resulted in the identification of the site as "Cussetuh" and defined the Lawson Field phase (Willey 1938). These and later excavations by Willey and crewmembers, which were determined to be located to the west and south, but close to the SERDP project geophysical surveys of 2002, included five burials and at least two structures evident from post hole patterns (Foster 2005). Additional burials were accidentally discovered in the 1970s, but they were farther away from the geophysical survey area. Later, in 1997, New South Associates set out to determine the boundaries of the site and found a single component site north of the geophysical survey area. Finally, in 2001 Panamerican Consultants, Inc. conducted extensive excavations north of the geophysical survey area and discovered 348 features, including five burials, a great many post holes, and a handful of trash pits. Of primary importance to this project, they noted that few architectural remains were preserved and the feature density was quite low (Foster 2005).

From the foregoing it seems likely that burials and post holes are the principal types of archaeological features that can be expected to occur, with occasional trash or perhaps storage pits. Descriptions of Creek Indian Towns from other sources provide additional information on the types of features that might be found, including clay extraction pits where clay was mined for waddle and daub construction. In addition, typical houses are rectangular, and round council houses are also found in most Creek Towns, so post mold patterns showing these shapes might be expected. Since typical

Creek houses are built to last only a few years, it is more likely that post holes will take on a more random distribution as multiple structural patterns may be superimposed on top of one another, and as houses and entire villages shift along the riverside (Swan 1855).

Due to the elusive and sparse nature of the archaeological record at Kasita Town, it is not surprising that the geophysical data did not reveal clear cultural patterns (Section 4.6.4). Consequently, the original data layers and fusions were utilized to vectorize interpretations in GIS, but categories were defined based only on geophysical source, anomaly shape, and strength. In other words, cultural interpretations could not be made because few, if any, were clear. As Kasita was the last site subjected to fieldwork, these very broad and general anomaly assignments completed a trend that began at Army City with very specific anomaly types linked with particular archaeological categories to one where archaeological assignments were not made.

Table 5.19. Anomaly classes for Kasita Town vector interpretation. *Description of 15 defined anomaly classes at Pueblo Escondido.*

Class	Name	Line / Polygon	Description
1	Mag lines	L	linear anomalies high in magnetometry data
2	mag lines, subtle	L	same as above, but more subtle and questionable
3	Mag polys	P	non-linear polygonal anomalies high in magnetometry data, not dipolar, mostly > 1 m diameter
4	GPR lines	L	linear anomalies high in GPR data, 0-80 cm depth (some are 1.4 - 2 m deep)
5	GPR polys	P	non-linear polygonal anomalies high in GPR data, 0-80 cm depth, mostly > 1 m diameter
6	Res lines	L	linear anomalies high in Resistance data,
7	Res polys, high	P	areas of high resistance, ranging from ~ 3 - 20 meters diameter
8	Res polys, low	P	areas of low resistance, ranging from ~ 3 - 20 meters diameter
9	Mag ditch, robust	L	broad linear positive anomalies in magnetometry, characteristic of ditch or related feature
10	Mag ditch, subtle	L	same as above, but more subtle and questionable
11	GPR ditch	L	broad linear anomalies in GPR data, approx. 1.2 m deep and deeper, characteristic of ditch or related feature
12	Res ditch, high	L	broad linear positive anomalies in Resistance data, characteristic of ditch or related feature
13	Res ditch, low	L	broad linear negative anomalies in Resistance data, characteristic of ditch or related feature

Abbreviations: *GPR* (Ground-penetrating radar), *Mag* (magnetic gradiometry), *Res* (electrical resistance)

Anomalies for each geophysical data type at Kasita Town were identified and categorized as lines and polygons, and some were divided between subtle and robust versions. These categories are summarized in Table 5.19 along with brief descriptions. Categories 9-13 were used to characterize the ditch-like features so clearly visible in every data set (e.g., Figure 4.46). Although these anomalies were collectively referred to as “ditches,” it was not assumed that they were cultural or associated with Kasita Town, as they are at least equally likely to be geological.

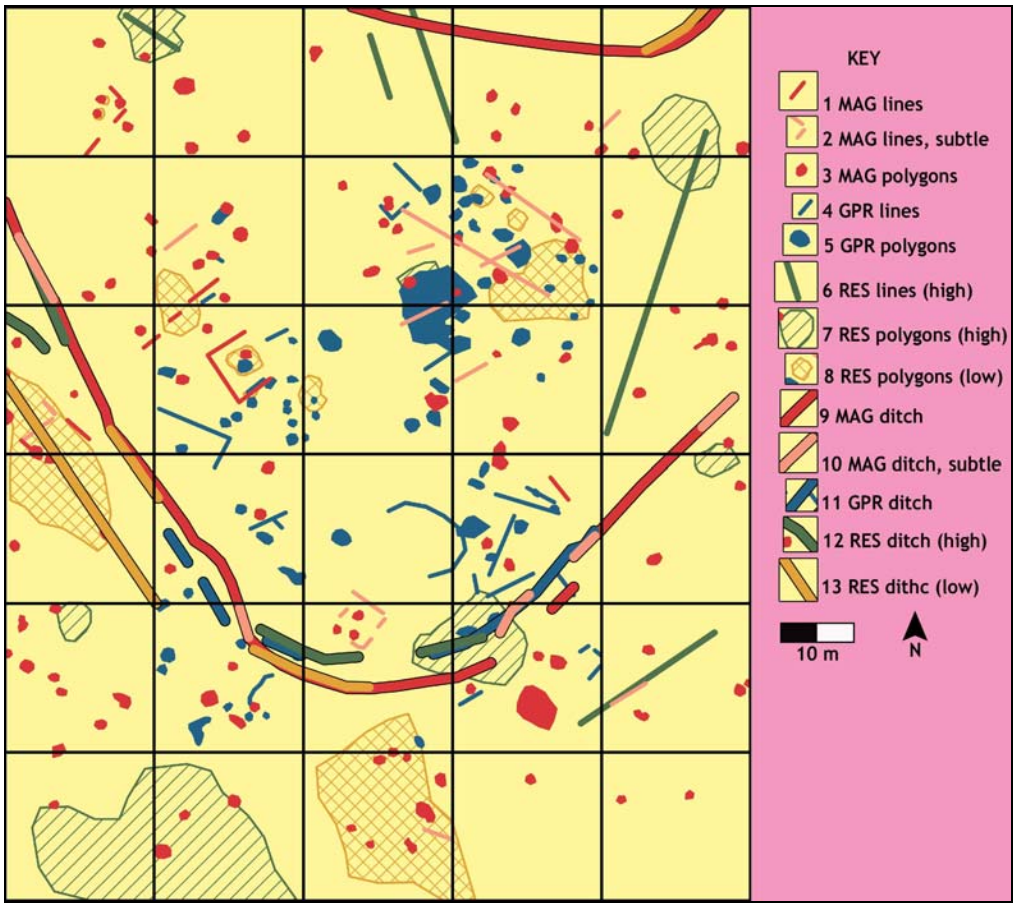


Figure 5.47. *Vector interpretation for Kasita Town. Linear Anomalies are exaggerated in thickness to aid visibility.*

The spatial distribution of the defined anomalies at Kasita Town is given in Figure 5.47. A region of relatively few anomalies running approximately down the approximate center of the study area is not an area of few geophysical anomalies. Rather, it is the location of a known metal pipeline that is strongly visible in the geophysical data. Anomalies close to this zone were not identified for testing because they could be disturbed or may have directly originated from the pipe's installation. For the remaining area, the majority of anomalies are not identifiable by shape so their possible identifications can only be speculated. The small to medium sized polygonal anomalies in all data sets could be burials, trash or storage pits, or other small features, including natural ones like badger dens or the sites of former trees or bushes. Larger polygonal anomalies are perhaps more likely to be clay extraction pits or dense concentrations of smaller features. Linear anomalies, besides the broad ditch-like anomalies, could possibly indicate structure walls or closely spaced alignments of post holes. The very large polygons, particularly those showing high and low resistance areas (Figure 4.46), are more likely to be geological in origin, but they could also be associated with Kasita Town agricultural land use or large midden deposits. The ditch like feature could indicate a fortification system, but this is unprecedented in Creek villages. It is therefore more likely to be a paleochannel or recent disturbance.

5.13. ISSUES IN FIELD VALIDATION OF ANOMALY PREDICTIONS

Kenneth L. Kvamme & Eileen G. Ernenwein, University of Arkansas

The two primary goals of the SERDP Project were to (1) develop means to combine or integrate multi-sensor information to better reveal subsurface archaeological content, and (2) validate predictions derived from the integrated data through a program of archaeological field excavation. The following sections examine the second primary goal of the project. Field validation of archaeological remote sensing anomalies is not only a complex topic, but a very specialized one that requires knowledge of (1) the methods and theories behind remote sensing, (2) the effects of soils and environment on remote sensing results, and (3) the nature of the archaeological record with specific insights into cultural forms that might be anticipated at a particular archaeological site associated within a specific range of time. The field validation phases of this project were one of its largest undertakings. They included the establishment of field sampling designs for each of four project sites, fieldwork at those sites with numerous excavations, descriptions and reports of archaeological findings, assessments of correspondences between remote sensing predictions and what was actually found, and final modifications to site interpretations based on field findings. Each of these topics is addressed in a separate following section.

5.13.1. Issues in Anomaly Source Validation and "Ground Truthing"

It is prudent to first examine exactly what validation or confirmation of a remote sensing prediction might mean. Frequently, as explained in Section 4.5.3, the sources of many remote sensing anomalies can be identified by their apparent form or distribution—a rectangular outline of the right size likely represents the foundation of a house, for example. Sometimes these patterns are so clear there can be no doubt about an anomaly's identity. In this SERDP project many such instances of this phenomenon occurred, as illustrated in Figure 5.49a-c. Some conservative archaeologists might still demand confirmation of the identity of such anomalies through independent evidence, however. On the other hand, in some sites the sources of many or even most anomalies may not express a form characteristic of a recognizable archaeological entity. In these cases archaeological excavation is the most used and most reliable means for identifying an anomaly's source. In other words, the archaeologist digs a hole at the anomaly's locus and attempts to identify some physical property or form that explains it. In most cases a physical source can be seen in the ground. Excavation has long been the central method of archaeology and observations made in the soil make up the bulk of the knowledge- and experience-base of that discipline. It remains the standard by which geophysical results are judged and forms the rationale behind seeking validation by excavation.

But excavation is not the only method for identifying an anomaly's source. Hargrave (2006) notes that a wide variety of information can play a role, including aerial and other photographs, early maps and archival records, geological maps, and anecdotal information from local informants. Excavation, however, remains the single most important technique for validating the results of near surface geophysical and remote sensing surveys of archaeological sites. Despite these varied means for validation, Hargrave (2006) and others prefer the term "ground truthing" to refer to this process.

In Figure 5.49d-e, excavation does indeed confirm a building footing and a concrete wall at Army City, but not the presence of an obvious pithouse at Pueblo Escondido (Figure 5.49c,f). Does this mean the pithouse does not exist? Any reasonable viewer would say "no"—the remote sensing evidence is simply too strong and there is no natural process that can explain the obvious structures. Also at this site and at Silver Bluff Plantation, strong lineations traversed the study areas in two directions, meeting at right angles (Sections 7.6.2-7.6.3). They are strongly suggestive of cultural constructions, yet no compelling evidence was observed in the excavations to explain their sources.

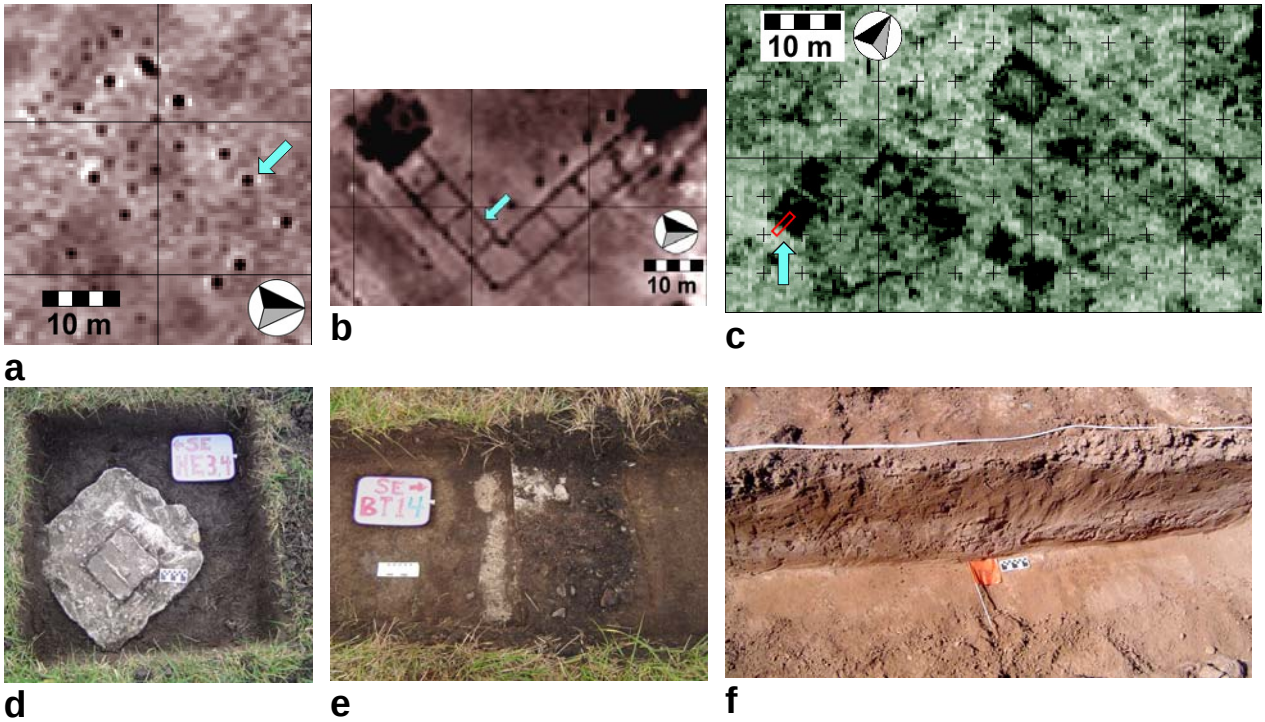


Figure 5.49. *Anomalies illustrating clear and unambiguous archaeological features indicated by geophysics and corresponding archaeological manifestations. a) Pattern of building footings in Army City resistivity data. b) Walls defining rooms of the Hippodrome in Army City resistivity data. c) Walls defining a line of adjacent pit houses in GPR data from Pueblo Escondido. d) Excavated building footing at Army City (at arrow in* a*) in 1 x 1 m hand excavation. e) Excavated concrete wall of Hippodrome at Army City (at arrow in* b*)(trench is 75 cm wide). f) 2.2 m segment of excavation profile through obvious wall of pithouse (at arrow in* c*) lacking observable indications of its presence.*

Some archaeologists might contest whether a predicted archaeological feature *invisible* in an excavation can be considered to exist when there is nothing in the ground for study (by traditional methods). Yet, geophysical instruments generally measure properties that often cannot be detected by human senses. A house floor, even if it has been heavily eroded or weathered, might be detected by geophysical instrumentation because the compacted soil might leave a "signature" in the shape of the house. That same floor might be undetectable with a trowel in excavation, however, but that does not mean it does not exist. It may indeed exist, albeit in a heavily weathered or secondary form. If remote sensing surveys are accepted as valid way of measuring the physical

properties of the ground, then anomalies showing probable cultural patterns should be accepted as primary evidence of archaeological features.

Remote sensing specialists take it as fact that, barring errors in instrumentation or data processing, an indicated anomaly is real and *must* have a physical cause. Consequently, there must be a source in the ground for *every* anomaly. The evidence of Figure 5.49f, in particular, illustrates that excavation is not infallible in providing conclusive answers to remote sensing questions. In other words, the tools commonly available to the archaeologist—the sense of sight that might reveal color differences between soil types or the sense of touch that might indicate hardness or texture differences—may not detect a physical change that explains an anomaly. This suggests that validation by excavations limited only to these senses may not be enough to reveal all anomaly sources. What is seen in an excavation may have no connection with geophysical measurements of properties that are often invisible to the human eye. If the goal of test excavations is to identify the source of a geophysical anomaly, then additional geophysical tools may be necessary for validation. Laboratory tests of soil samples taken from excavations or in-field tests with instrumentation may be necessary to indicate physical properties responsible for anomalies that have no observable source. A more thorough evaluation may be achieved through use of soil tests for relative changes in sediment size, organic material, density, moisture content, and other physical properties. Additional, targeted geophysical measurements of magnetic susceptibility and conductivity could also be employed to shed light on anomalies during test excavations. Hand-held sensors and lab-facilitated measurements of soil samples could provide valuable and more conclusive information to aid evaluations.

The primary issue in validating remote sensing findings is not whether an anomaly has a physical source, but whether its *interpretation* is correct. At the most basic level is the question of whether an anomaly's source derives from natural or cultural causes. Beyond this lies more specific identifications: whether an anomaly claimed to represent a house is associated with a real a house below the surface or whether one classified as a wall actually represents a buried wall. Excavation has shortcomings in validating these kinds of interpretations as well. Often not enough is known about the content of the archaeological record or its formation processes for a solid understanding of physical forms that might represent an anomaly's source. What does the surface of an unpaved historic street or a rut in that street look like archaeologically? How should informal earthen gutters bordering those same streets appear in the soil? These were pressing questions at Army City but answers were not always satisfying or clear, partially because the limited archaeological knowledge base cannot always provide definitive answers. Although physical evidence of a *difference* might occur at the source of an anomaly, it might not be clear to the archaeologist what kind of cultural feature (or indeed whether it is a cultural feature as opposed to a natural one) is represented owing to the many vagaries associated with archaeological finds and states of preservation. This problem is exacerbated by the typically small holes that archaeologists must employ as cost-cutting measures. A 35 x 35 cm shovel test pit (STP) placed over an anomaly suggested to represent a "wall" might reveal a dense burned layer containing charcoal. An explanation for the anomaly is seen, but does it represent a wall that might have burned? Only a larger excavation that shows a sufficient length of its linear form and demonstrates it is not merely a burned log can confirm its identity. The foregoing errors

indicate that "truth" may not always be consistently found in the ground, which has led some researchers to feel uncomfortable with the term "ground-truthing." Other specialists nevertheless use that term and it is employed in some of the following sections to refer to general anomaly validation efforts.

Many of the previous difficulties are made worse because few archaeologists are knowledgeable about remote sensing (it is not taught in most university archaeology programs). They are trained specifically to understand subsurface soil changes, archaeological deposits, and identify such archaeological features as the remains of architecture and other human constructions. Yet, although they are trained to recognize *archaeological* causes of indicated anomalies (e.g., a prehistoric storage pit or house floor) they may not fully appreciate other possible causes, such as the filled-in badger den, tree root, or natural soil change. This is probably truer of those archaeologists who actually perform most excavations because they tend to come from the lower end of the intellectual ladder (i.e., young or recent students holding Bachelor of Arts degrees).

5.13.2. Field Assessment Tactics

So concerned was the SERDP team with these issues that they performed independent evaluations of every sampled anomaly at each of the study sites. As described in Section 5.15, each excavation was evaluated independently by a team of archaeologists *and* by the SERDP team. The latter was able to spend much more time evaluating the excavations than the archaeologists who were also responsible for many other tasks, including excavation, screening, mapping, artifact packaging, and documentation of results. The SERDP team was able to cut back or re-face trench walls and floors as required for a clearer view of soil changes and other variations, use soil augers to check depths and extents of visible features, and in some cases employ a hand-held magnetic susceptibility meter to better understand deposits. By these means they were frequently able to zero-in on the physical sources that explained many of the observed anomalies.

At the same time, it is emphasized that the archaeological teams were absolutely essential to the validation process, meeting two important needs. First, each of the selected sites were eligible for nomination to the National Register of Historic Places, requiring professional excavations, curation of recovered artifacts, and reports submitted to relevant agencies. The SERDP team did not have time for these tasks and lacked local expertise in some areas. More importantly, most assessments of remote sensing in archaeological projects are and will be made by archaeologists. It is therefore important to realize the kinds of performance and errors they obtain and the mistakes they realize in remote sensing evaluations.

A few examples illustrate some of the foregoing issues, but also reveal more basic difficulties stemming from terminology. In the interpretation of anomalies the SERDP team initially categorized them into likely archaeological types (Section 5.12). At Army City, the first site subjected to field validation, some anomalies were inferred to represent walls, others floors, pipelines, or street gutters, for example. These very simple labels nevertheless led to many difficulties. One linear anomaly revealed primarily by GPR was classified as a possible wall about 4 m long (Figure 5.50a), but excavations showed it to be a smashed ceramic sewer pipe near the surface that was probably dropped from a truck during the town's dismantling (Figure 5.50b). The remote sensing team was pleased

because an obvious source for that anomaly was readily found; the archaeologists, however, saw and recorded this find as an error in the remote sensing predictions because a "pipe" is obviously not a "wall" (Kresja and McDowell 2005:40).

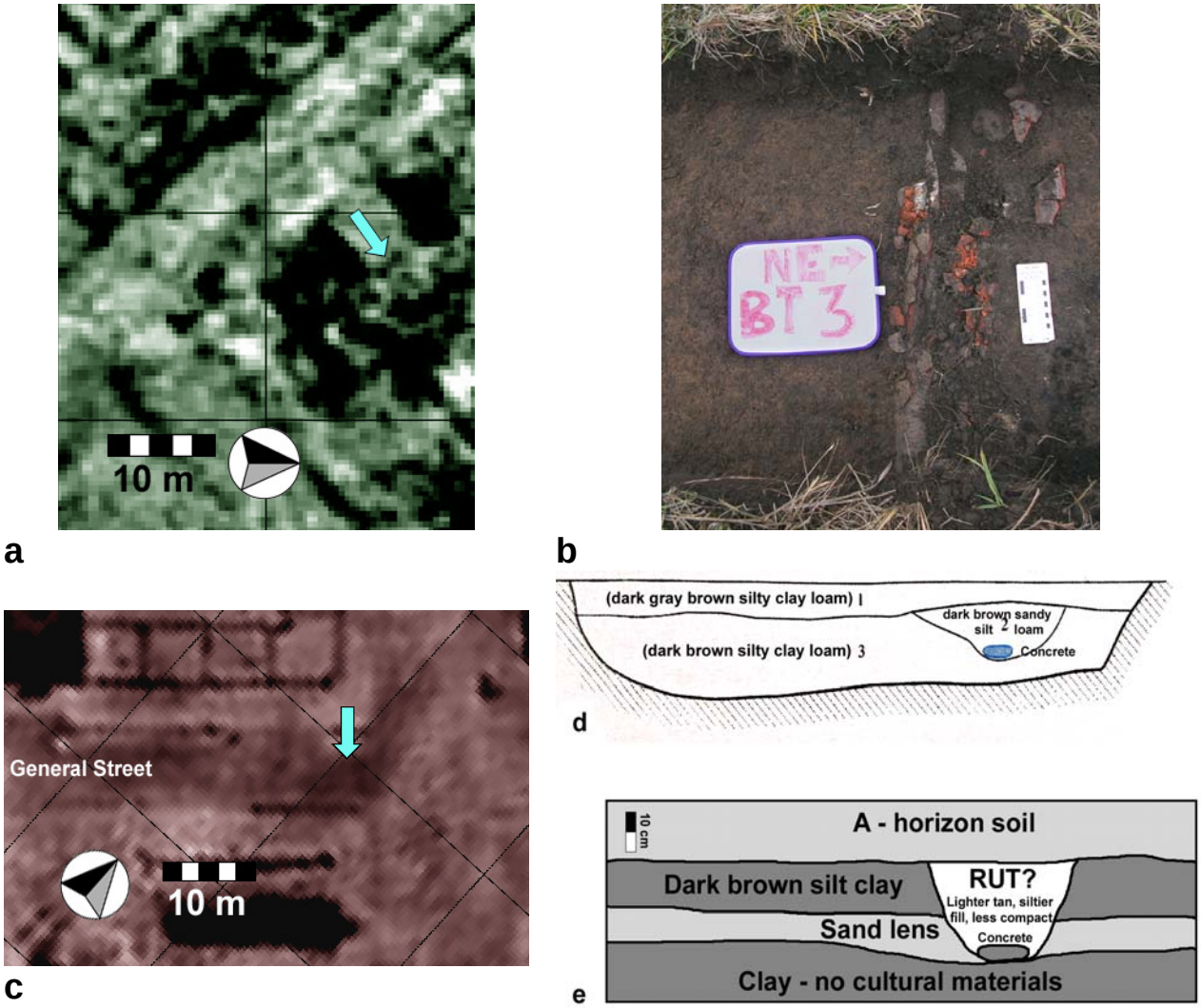


Figure 5.50. *Evidence for anomaly identifications at Army City. a) Portion of GPR survey data showing linear anomaly interpreted as a "wall" (arrow). b) Excavation at the locus of the wall anomaly revealing a broken ceramic sewer pipe (trench is 75 cm wide). c) Data showing resistant street anomaly (arrow). d) Profile of Trench BT 25at center of anomaly in* c *showing stratigraphy mapped by the archaeological team that found no evidence of a hard-packed street surface (after Kresja and McDowell 2005:83). e) Profile of the same trench by the SERDP team that indicates a layer of sand that explains the anomaly.*

A more complex example also occurs at Army City where a segment of one street exhibited very high resistivity, causing the remote sensing team to come up with several hypotheses that might explain this anomaly (Figure 5.50c). It was possibly surfaced, baked, hard-packed, or perhaps strewn with gravel or sand—all reasonable circumstances that could yield high resistivity. Unfortunately, in the confined space of the field form

used by the archaeologists for their assessments (see Section 5.15), a label listing only "street-hard/packed/baked/surfaced" was employed. The archaeological team observed the profile mapped in Figure 5.50d and, although they noted a thin layer of sand partially exposed on the trench's surface (Kresja and McDowell 2005:82), because it was not hard-packed that anomaly's predicted type was classified as incorrect. On the other hand, the SERDP team, with more time and intent on locating a physical explanation, used shovels and trowels to more intensively scrape and examine the trench walls. In so doing they defined a different stratigraphic sequence that included a subtle layer composed largely of sand primarily revealed as a texture difference from 8-22 cm in thickness (Figure 5.50e). As sand is one of the most resistant of sediments, this allowed their reliability assessment for this anomaly's predicted type as "correct" with "high confidence," an enormous difference from the other evaluation.

These kinds of difficulties forced the SERDP team to move away from specific archaeological classes and terms to more general ones in subsequently evaluated sites. Terms like "robust linear feature" or "subtle area feature" were employed in order that terminological uses not confound the intent of the exercise—to ascertain whether remote sensing predictions can be associated with reasonable archaeological explanations.

5.13.3. Research Plan

The following sections describe the field validation efforts of the SERDP Project. Sampling designs are first established for each site that define which anomalies are to be evaluated. After a discussion of our field methods, the results of the field evaluations are presented from each site from two perspectives, that of the SERDP team and that of the archaeologists who worked on each site. These evaluations include assessments of the accuracies of the remote sensing predictions. Finally, a few of the insights learned from each evaluation project that augment knowledge of the sites are summarized.

5.14. SAMPLING DESIGNS FOR ARCHAEOLOGICAL FIELDWORK

Kenneth L. Kvamme & Eileen G. Ernenwein, University of Arkansas

Consideration of a program of field sampling to test and validate the predicted anomaly classifications of Section 5.12 began with confrontation of a fundamental question: whether excavations should (1) be conducted at the most "interesting" or "promising" anomalies in a judgmental form of sampling, or (2) employ a program of random sampling. Both offered advantages. The former could target most "interesting" or "promising" anomalies that might indicate a great deal about the sites and allow exposure of archaeological features of high significance. Each remote sensing data set revealed numerous anomalies that were not only highly interpretable, but also archaeologically interesting and sometimes puzzling (Section 4.6). How tempting it was to simply target these anomalies, but in so doing primarily robust, clearest, and easiest-to-interpret anomalies would have been selected and "stacked the deck" towards high rates of predictive success as well as a biased view of the sites' contents toward robust and interpretable elements. A random sampling program, on the other hand, could yield an unbiased view of the performance of the remote sensing program, a clearer view of site content and anomaly correspondences in general, and defensible measures of performance accuracy.

The decision became an easy one through consideration of the project's goals—methods for processing and integrating multi-sensor data that enhance the potential for detecting *and characterizing* subsurface features. It would not be sufficient to simply ascertain whether particular anomalies were associated with cultural features. The project's objectives required assignment of anomalies to interpreted archaeological categories (as discussed in Section 5.12, the project began with anomalies linked to specific archaeological classes and ended with more general anomaly types, but even they were linked with likely or possible archaeological causes). It was therefore necessary for the archaeological testing phase of the project to determine whether assignments of anomalies to such categories were reliable. This need required the investigation of representative samples of the anomalies assigned to each archaeological type. Simply focusing excavations on the "best" examples of each category would not permit determination of the effectiveness of the project's data processing, integration, and interpretive protocols. In other words, these protocols might allow the detection and accurate interpretation of only the most obvious anomalies. Evaluating the worthiness of the project's methods also requires attempts to correctly interpret amorphous or otherwise ambiguous anomalies.

All interpretations of remote sensing anomalies must be tentative until their existence and specific nature can be further investigated by subsurface archaeological excavation or other means (Section 5.13). Since each anomaly type defined by this project includes from several to many individual examples of similar size, shape, or geophysical signature (Section 5.12), the archaeological testing of a sample allows refinement of interpretations across the entire survey area. The process of sampling ultimately makes it possible to make reasonably confident interpretations about an entire site (at least within the confines of the geophysical survey boundary) by excavating a very small but targeted percentage of the total area.

Ultimately, stratified random sampling designs were formulated for each site that included anomalies assigned to each functional or descriptive archaeological category. Construction of the sampling programs had to confront several issues. Test excavations conducted at large samples of anomalies were desirable and an obvious goal. Large samples would potentially allow reliable performance statistics and confidence intervals to be computed for each class of anomalies. Yet, from a pragmatic standpoint, large samples require great archaeological effort, and archaeological fieldwork is expensive and destructive. Limited budgets dictated relatively small sample sizes. Moreover, because the project sites were potentially eligible for the National Register (Section 3.3), destructive impacts, which include archaeological excavations, had to be minimized. Compromises were therefore reached between budgets, numbers of samples, and the nature of the excavation impacts. The outcome was rather limited sample sizes from a statistical standpoint. Low numbers prevailed in each of the site sampling designs. With such small samples it is not possible to make meaningful statistical evaluations of performance for individual classes of anomalies (Section 5.16). Assessments of aggregate classes and overall performances of the remote sensing classifications could be made, however.

5.14.1. Types of Excavation Samples

Several forms of test excavation were possible in the SERDP project sites. Each varied in its "costs, " amount of labor required (and funding), destructiveness to the site, and area of exposure for field crews to evaluate (which relates to the "information content" each yields). From least to most invasive they include the following.

1. *Soil cores* are made with a one-inch (2.54 cm) Oakfield corer, which is pressed into the ground to typical depths of one-half to one-meter. In the soil plugs brought to the surface one can assess the nature of soils and sediments, whether they may be "cultural" in nature, and even witness artifacts and evidence of anthropogenic activity (e.g., charcoal, ash, bone). In rocky sites or dense clays it may not be possible to insert core holes into the ground.

2. *Shovel test pits (STP)* are small holes excavated with a shovel or spade, usually measuring about 40 x 40 cm in area and commonly excavated to a depth of 40-60 cm. They offer some visibility of the subsurface and expose profiles in which stratigraphy can be assessed.

Soil cores and STP can be placed singly, or in a "paired" format where one is placed inside a targeted anomaly and a second is located immediately outside of it for comparative data.

3. *Shovel scrapes* are utilized in the Southwest instead of STP. In this form of exposure a shovel is employed to scrape or skim the soil a few centimeters at a time in order to better witness archaeological features. They are typically the width of a shovel (about a quarter-meter) by 1-2 m long.

4. *Hand excavation units* represent formal excavations in the classic textbook sense. Because of the effort and time they require, they are typically small in size measuring perhaps from .5 x 1 m up to 1 x 2 m. They are dug by hand with a combination of shovels and trowels in natural levels following stratigraphy or, when absent, in arbitrary levels typically 10 cm in thickness. They yield excellent floor plans, stratigraphic data, and high information content in general. Because of their expense and the large amounts of time required to excavate them, they are used sparingly and placed to clarify targets of high interest.

5. *Linear mechanized excavations* usually average just under a meter wide, the typical width of a backhoe's bucket. They are performed with a toothless bucket and require an expert and experienced backhoe operator. Excavations are made with a scraping motion where perhaps 3-6 cm of deposit is removed at a time. During this process archaeologists closely monitor the removal looking for telltale signs of archaeological deposits, features, or artifacts. Periodically, the scraping is halted while archaeologists scramble to scrape with shovels or trowels to clarify floors and profiles and investigate suspicious signs. Resulting trenches may be 2-5 m long and perhaps 25 cm -1 m deep (depending on the site and nature of the

target). These units are most informative for anomaly assessment because they expose large areas and profiles. They are typically placed to crosscut a linear anomaly, such as one classified as a "wall," in order that the feature can be seen in plan and profile and the excavation exposes areas on both sides (perhaps showing regions "inside" and "outside" of a room, for example).

This ranking of excavation methods also relates more or less to expense and information content. Soil cores are least expensive, but least informative. STP give much more data but require more effort and therefore expense. Hand excavations can be extremely informative and well document the subsurface, but are most expensive and time consuming. Mechanized excavations are relatively efficient and quickly achieved. They are hugely invasive, but the large exposures of plans and profiles yield maximum information potential for recognition of cultural features and assessment of their correspondences with anomalies. These methods are illustrated in Section 5.15 where field methods employed by this project are fully described.

As in foregoing sections, approaches taken at Army City illustrate how the foregoing issues were weighed and how sampling methodologies were constructed. Succeeding sections then overview sampling at the remaining project sites.

5.14.2 Sampling Army City

Establishing a sampling design for anomaly testing is a two-stage process. The first step requires definition of anomaly types represented in the multidimensional data (performed in Section 5.12), and the second assigns elements to be sampled, presented here.

5.14.2.1. Types and numbers of samples (excavations)

In creating a sampling design several considerations were paramount. One was as large a sample size as possible within the confines of limited budget. Another was sampling of each indicated anomaly type. Budget limitations suggested the possibility of excavating 200 linear meters of trenches by mechanized equipment, 12 hand-dug excavation units measuring 1 x .5 m, and 50 STP measuring about 40 x 40 cm, all to a depth of 40-50 cm on average (well into the site's target depth). Soil cores were ruled out because of the high clay content of the site. These excavation quantities were broken down as follows.

- *Linear mechanized excavations*. Averaging a meter wide, a length of three meters was deemed sufficient to crosscut linear features to make their presence clear against the normal background. Sixty-six trenches of 1 x 3 m were planned.
- *Hand excavation units*. Twelve such units were planned, each measuring 1 x .5 m. Six were placed as part of the initial sampling design. Some of these were to allow a quality photograph and mapping of anomaly types. The remaining six were placed judgmentally as part of a "staged" approach to testing (Section 5.15.1), for example, over anomalies in trenches that required further clarification.
- *Shovel test pits*. Fifty STP were planned. Eighteen were "single" STP designed to be placed over anomalies predicted to be unambiguous and easily located (e.g., over likely metal targets). Others were used to test the "anomaly absent" class.

Sixteen "paired" STP were also placed, particularly over building footers or floors. The total number of STP samples is therefore 18+16=34 (with the remaining 16 forming the pairs).

- *Total number of samples*. This sampling program allowed 66 mechanized trenches, six hand-excavated units, and 34 STP, for a total of 106 samples. An additional six hand-excavated units and 16 STP were set aside to give supportive evidence and additional insights.

5.14.2.2. Sampling design

Army City's defined anomaly types (Table 5.16) were stratified into four classes, with the background forming a fifth class. To put most effort in the testing and examination of anomalies, it was decided to place only 10% of available test samples in the background class, with the remaining 90% distributed equally among the four strata describing archaeological anomalies (at 22.5% each). The association of each anomaly type with stratum as well as population and sample sizes is given in Table 5.20. Sampling was accomplished in the following manner.

- *Background class*. A three-meter buffer was generated around the Boolean Union of all anomalies indicated in Figure 5.44e. Background points for STP placement were selected from the remaining area by random coordinates (i.e., from regions greater than 3 m from any defined anomaly). The buffer insured a distance sufficiently far away from defined anomalies to avoid contamination by them. It emphasized that areas of mechanized trench excavations centered on anomalies that extend past them also provide ample evidence of anomaly absence, as do the second of paired STP.

- *Anomaly classes*. For each anomaly class of Table 5.16 the number of distinct polygons or arcs, depending on data type (Figure 5.44e), was counted. Rare classes (chimney [Class 1] and street surface [Class 10]) were forced to receive two samples each. Within each stratum the number of samples was taken in proportion to the number of distinct polygons or arcs in that class. Each distinct element was assigned a number and elements were sampled by random selection. The locus sampled in each element is generally near its center, except trenches placed in "floors" were it is desirable to intersect an edge or where many different features converge, where samples were moved away from the center. In some cases where random selection caused adjacent anomalies to be selected, another close-by anomaly of the same type was judgmentally selected.

Table 5.20. *Anomaly types, strata, population and sample sizes for the Army City excavation program.*

Code	Anomaly type	Stratum	Population size (# of polygons or arcs in type)	Test unit types				Total samples
				T	EX	S-STP	P-STP	
1	*Chimney base*	M	1		2			2 (adjacent)
2	*Footer*	M	39		2		6	8
3	*Isolated metal*	M	17			3		3
4	*Pipe*	M	24		1	4		5
5	*Debris*	M	32	5				5
			Totals:	**5**	**5**	**7**	**6**	**23**
6	*Robust wall*	W	81	10				10
7	*Subtle wall*	W	106	14				14
			Totals:	**24**				**24**
8	*Robust floor*	F	19	7			5	12
9	*Subtle floor*	F	14	5			5	10
10	*Street surface*	F	1	1	1			2
			Totals:	**13**	**1**		**10**	**24**
11	*Robust street gutter*	S	22	7				7
12	*Subtle street gutter*	S	23	7				7
13	*Street rut*	S	7	2				2
14	*Robust alley gutter*	S	10	3				3
15	*Subtle alley gutter*	S	9	3				3
16	*Subtle sidewalk edge*	S	3	2				2
			Totals:	**24**				**24**
17	*Background*	B	∞			11		11
			Totals:	**66**	**6**	**18**	**16**	**106**

Stratum Abbreviations: M=Miscellaneous; W=Wall/linear; F=Floor/surface; S=street feature; B=background. **Test Unit Type Abbreviations:** T=mechanized 1x3m trench; EX=1x.5m hand excavation; S-STP=single STP; P-STP=paired STP.

The spatial distribution of the sampling design, by test unit type, is given in Figure 5.51a. In addition to these 106 units, an additional 16 STP were placed outside of anomalies to offer "paired" data to contrast with anomalies. Moreover, six additional 1 x .5 m hand excavated trenches were judgmentally placed to clarify certain anomalies. To aid in fieldwork, the 106 excavation units were also superimposed over imagery showing the geophysical results, to indicate to fieldworkers the nature of the primary evidence and facilitate targeting of anomalies (e.g., whether they might derive from a metal pipe [Figure 5.51b] or from a concrete wall [Figure 5.51c].

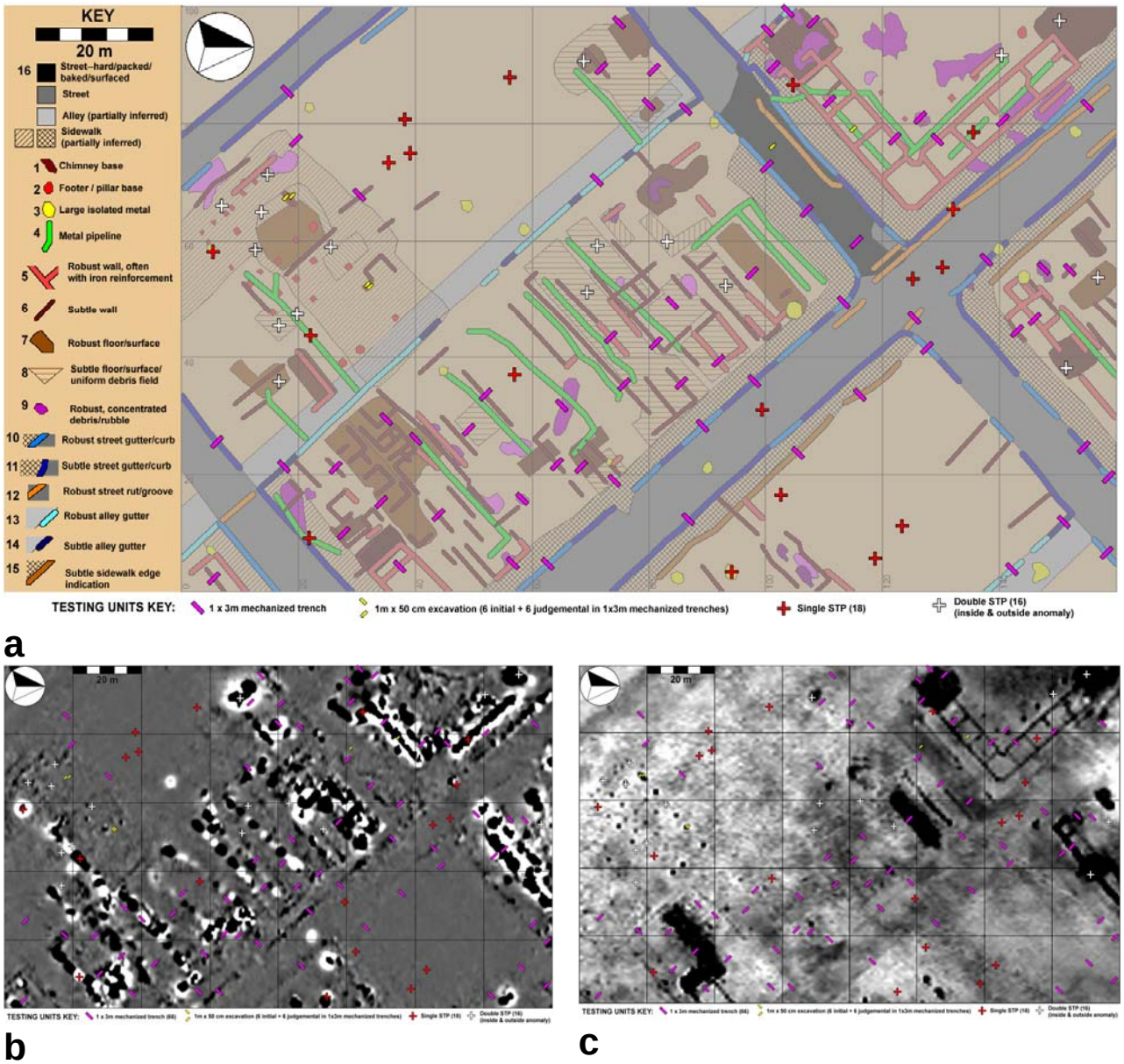


Figure 5.51. *Archaeological sampling design and distribution of excavation types at Army City a) overlaid on vector interpretations, b) in the magnetic gradiometry data, and c) on the electrical resistivity data.*

5.14.3. Sampling Design for Pueblo Escondido

The geophysical data reveal remarkable detail about the spatial layout of Pueblo Escondido (Figures 7.38-7.44). A vector interpretation of these data was created that includes 15 anomaly types (Table 5.17), and communicates the spatial details of the site (Figure 5.45). The first step in planning a sampling strategy for Pueblo Escondido involved determining the extent of excavation that would adequately evaluate the geophysical data while causing minimal disturbance to the site and staying within the confines of a limited budget. It was not known what kinds of units would most efficiently and accurately allow the evaluation of each anomaly type, so a variety of different

excavation units were utilized. These factors culminated in a plan to excavate 100 linear meters of trenches by mechanized equipment, 60 shovel-scrapes measuring .25 x 2 m, 20 hand excavation units measuring .5 x .5 m, and 80 augur or core holes. The 100 linear meters of trenches was distributed among nine units measuring three meters in length, and 12 trenches measuring two meters in length. Shovel scrapes are shovel-width sized trenches recommended by the local contracting archaeologists. Core holes were utilized to test the background class, and also to test features predicted to be unambiguous and easily identified (such as hearths). Many core holes were set up as paired comparisons, where one core was placed within an anomaly and the other outside to provide contrast.

Although the primary goal of the excavations was to quantify the performance of the geophysical methods, a secondary goal was to better understand Pueblo Escondido through this rare excavation opportunity. A stratified random sampling procedure was designed to test each of the 15 defined anomaly types (Table 5.17) in proportion to their "population" numbers. Four strata were defined that grouped similar types of anomalies: (1) those within rooms or houses, (2) linear anomalies outside rooms or houses, (3) polygonal anomalies outside rooms or houses, and (4) the undisturbed background (the absence of anomalous indications). The background class is a vast area compared to the classified anomalies, and unlikely to contain many archaeological features. The background stratum was therefore tested by 10% of the excavation units, while the remaining 90% were distributed according to the number and of identified features in each stratum. Another component of the sampling design was the reservation of 18% of the test units (excluding auger tests) to be placed judgmentally during field work where needed based on findings. The distribution of anomaly types in each stratum is summarized in Table 5.21, which lists the number and types of test units placed in each.

The background class was primarily sampled by random placement of core holes. It was also sampled by careful placement of comparison cores in the paired sampling, and by those excavated areas extending beyond targeted features. The remaining three strata were sampled almost equally, with a slight shifting of the weights to the strata with more total anomalies and greater total area. Within each stratum rare anomaly types were forced to receive at least four samples, and the remaining were each assigned numbers and randomly selected. If the same anomaly was selected more than once a new random number was generated until an untested anomaly was selected. The exact locus of a test unit along an anomaly was judgmentally placed in order that it could be examined without complications from nearby anomalies. The distribution of excavation types is shown in Figure 5.52.

Table 5.21. *Anomaly types, strata, population and sample sizes for the Pueblo Escondido excavation program.*

Code	Anomaly type	Stratum	Population size (# of anomalies /, total area* in m^2)	Excavation unit types					Total samples
				M	EX	SS	P-C	S-C	
1	*Room block floor*	R	12 / 150.6	1	1	2	1		**5**
2	*Magnetic room block floor*	R	14 / 90.9	1	1	3	1		**6**
3	*Inferred room block floor*	R	20 / 637.7	1	1	1	1		**4**
4	*Outlying floor*	R	21 / 147.3	2	1	3	1		**7**
5	*Magnetic room block wall*	R	21 / 86.3	2	1	4	1		**8**
6	*Robust room block wall*	R	34 / 122.0	3	1	2	1		**7**
7	*Subtle room block wall*	R	17 / 51.2	1	1	2	1		**5**
8	*Interior room block feature*	R	16 / 16.2	1	1	2	1		**5**
9	*Circular magnetic interior room block feature*	R	27 / 32.6	0	1	2	1	2	**6**
			Totals: 182 / 1334.8	**12-1x2m**	**9**	**21**	**9**	**2**	**53 (35%)**
10	*Robust lineation*	L	72 / 255.8	6	1	11	1		**20**
11	*Subtle lineation*	L	31 / 103.2	3	1	5	1		**10**
			Totals: 103 / 359.0	**9-1x3m**	**2**	**16**	**2**		**29 (20%)**
12	*Small polygon systematically distributed*	P	180 / 336.5	8	1	10	10		**29**
13	*Small polygon randomly distributed*	P	93 / 83.8	4	1	6	7		**18**
14	*Circular magnetic exterior anomaly*	P	46 / 65.4	1	1	0	1	2	**5**
			Totals: 319 / 485.7	**13-1x2m**	**3**	**16**	**18**	**2**	**52 (35%)**
15	*Background*	B	∞ / 7820.5					18	**18 (10%)**
			Totals: 604 / 2179.5	**34**	**14**	**53**	**29**	**22**	**152**
18% of M, EX, SS reserved for "staged" sampling				**5-1x2m; 4-1x3m**	**6**	**7**			**22**
			Total units:	**43 99m**	**20**	**60**	**29 pairs**	**22**	**174**
							80 total		

Stratum Abbreviations: R=Room block; L=Major lineation; P=Point/polygon anomalies; B=background. **Excavation Unit Types:** M=mechanized 1x2m or 1x3m trench; EX=.5x.5m hand excavation; S-S =shovel scrape; P-C=paired core; S-C=single core. *Area for linear classes calculated as 1 linear meter = 1 square meter.

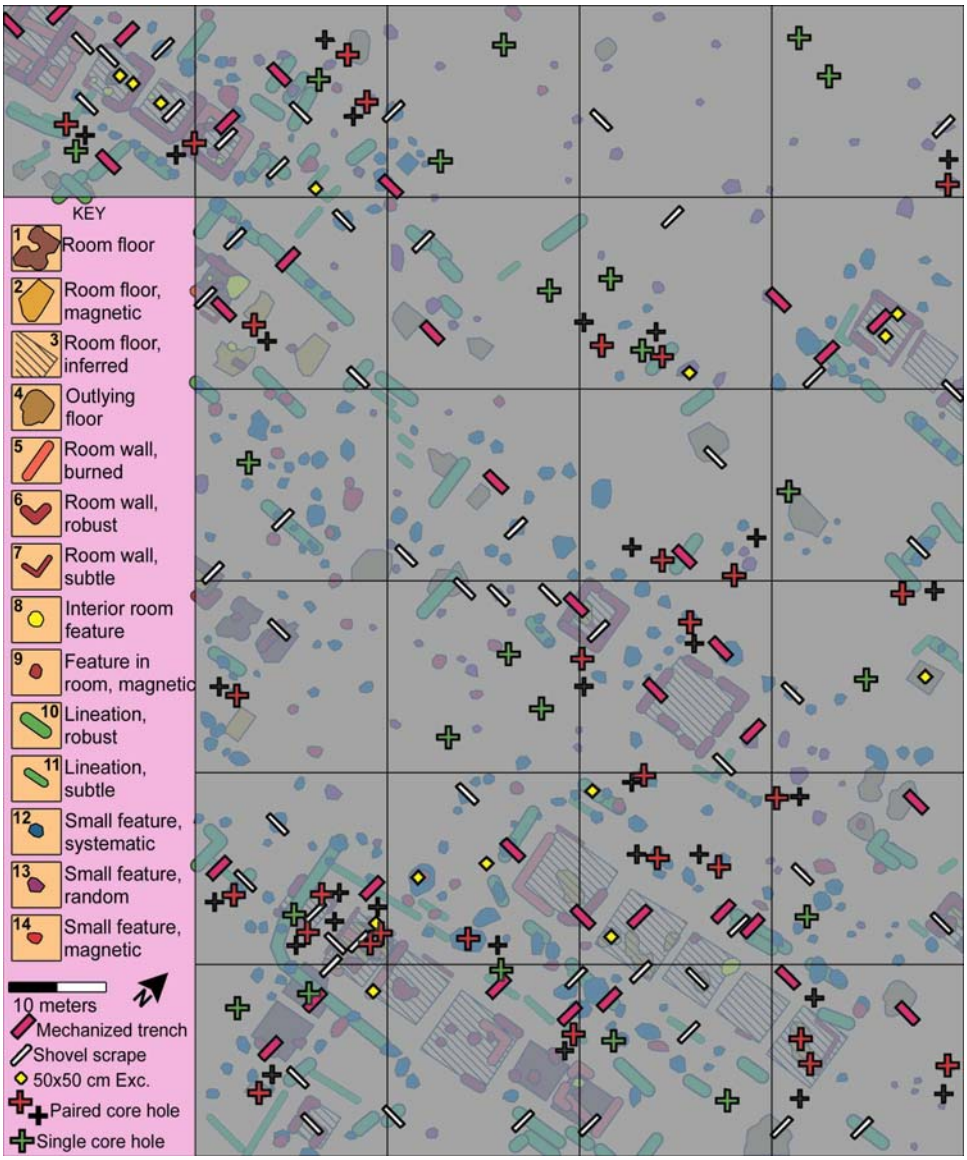


Figure 5.52. *Distribution of excavation units at Pueblo Escondido overlaid on the vector interpretation. Linear Anomalies and excavation unit symbols are exaggerated.*

5.14.4. Sampling Design for Silver Bluff

The sampling design established at Silver Bluff Plantation was based on a somewhat different class of interpreted anomalies. As indicated in Section 5.12.6, although clear cultural anomalies were revealed, lack of experience in the region's archaeology and the site suggested use of generalized anomaly types, so rather than labeling a robust linear GPR anomaly a "wall," it was simply referred to as a robust linear GPR anomaly (because it could represent a builder's trench or line of posts, for example). The "significant" anomalies represented three basic types: narrow and clear GPR anomalies, broad resistivity anomalies, magnetic anomalies primarily form ferrous metal or burning, plus a miscellaneous category. This indicated that the real "strata" corresponded primarily with the geophysical survey types. Whatever the case, given the nature of the site it was determined that only linear mechanized excavation (backhoe

trenches) could open enough space to allow fieldworkers to sort out the puzzles of the indicated anomalies.

Funding was available for 50 mechanized trenches measuring 1 x 4 m. Twenty STPs were employed to test and retrieve indicated ferrous metal artifacts (Table 5.22). Additionally, two 1 x .5 m excavations were placed over interesting linear GPR anomalies and three were reserved for judgmental placement over interesting features exposed by the backhoe. Oakfield core holes were to be liberally employed on a selective basis to check depth and extent of archaeological features exposed by mechanized excavations.

Linear anomalies classified in Figure 5.46 were numbered by individual arcs, as were magnetic point anomalies, and large polygons were broken up into "linear" segments (mimicking arcs) and numbered for many of the resistivity anomalies. These "population" numbers are given in Table 5.22. With the foregoing level of effort it was possible to sample approximately 20% of the anomalies. This was accomplished through simple random sampling for each class. Sample numbers for each anomaly class are also given in Table 5.22, and their spatial distribution is depicted in Figure 5.53.

Table 5.22. *Thirteen anomaly types at Silver Bluff Plantation. All anomaly classes treated as polygon data.*

Class	Name	Stratum	Population	Samples
1	*Historic iron or steel*	Mag	189	20*
2	*Subtle linear magnetic*	Mag	13	2
3	*Magnetic point in linear alignment*	Mag	10	2
4	*Linear negative resistance*	Res	20	5
5	*Linear positive resistance*	Res	51	10
6	*Area negative resistance*	Res	33	5
7	*Area positive resistance*	Res	3	1
8	*Subtle linear GPR*	GPR	94	18**
9	*Subtle negative linear GPR*	GPR	1	1
10	*Robust linear GPR*	GPR	17	4**
11	*Area GPR*	GPR	9	2
12	*Area robust GPR--deep*	GPR	1	1
13	*Lightening strike*	GPR	1	1
		Totals:	442	72

*STP; **one 1 x .5 m hand excavation in this category; all other 1 x 4 m backhoe trenches. Mag=magnetic gradiometry; Res=resistivity; GPR=ground-penetrating radar.

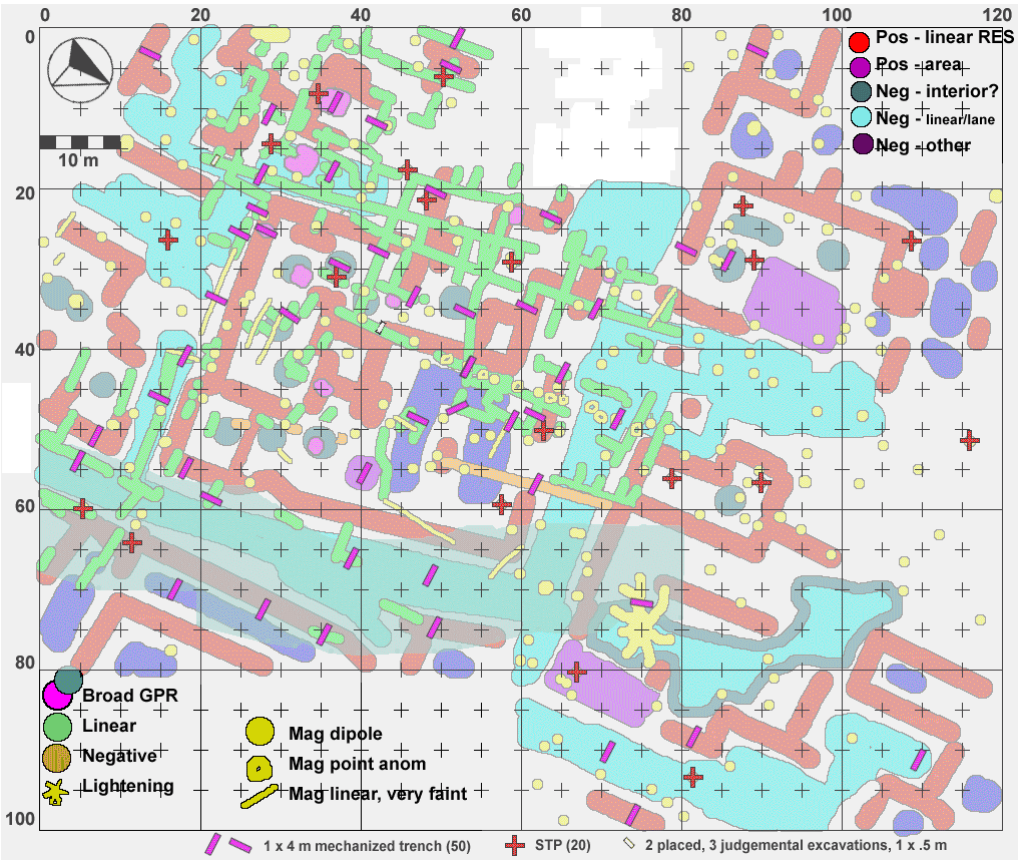


Figure 5.53. *Distribution of excavation units at Silver Bluff overlaid on the vector interpretation. Linear Anomalies and excavation unit symbols are exaggerated.*

5.14.5. Sampling Design for Kasita Town

The vector interpretation of Kasita Town (Figure 5.47) served as a starting point for planning test excavations. Based on the nature of the geophysical data and knowledge gained from previous excavations close to the survey area (Section 3.3.4.4), it was expected that many anomalies were naturally generated or resulted from recent Army-related disturbances rather than cultural activities related to the proto-historic occupation of the town. Two types of excavations were chosen to examine anomalies: long backhoe trenches (1 x 8 to 1 x 20 m) and 2 x 4 m strip blocks. The long backhoe trenches were designed to examine very large anomalies, particular broad areas of high and low electrical resistance and many of the "ditch" anomalies (Categories 9-13 of Table 5.19), which are several meters wide. The many other small and scattered anomalies (Categories 1-8 of Table 5.19) were tested with strip-blocks, with the idea that exposing a large surface area might reveal post hole and small archaeological feature patterns that would otherwise be more difficult to interpret in isolated trenches. Each stratum was tested according to the proportion of anomalies contained in that class compared to the total population of anomalies identified. With the exception of three backhoe trenches, anomaly categories 1-8 were tested with 2 x 4 m strip block units and categories 9-13 were tested with backhoe trenches of varying lengths as necessary to span the anomaly. Overall 30 strip blocks and 11 backhoe trenches were planned for a total of 360 square meters. The number and type of test units planned for each anomaly category and stratum are given in Table 5.23. The distribution of planned excavation units at Kasita Town is given in Figure 5.54.

Table 5.23. *Anomaly types, strata, population and sample sizes for the Kasita Town excavation program.*

Code	Name	Stratum	Population	SB	*T*
1	mag lines	M	11	4	
2	mag lines, subtle	M	16	3	1
3	mag polygons	M	111	7	
		Subtotals:	***138***	***14***	***1***
4	GPR lines	G	22	6	
5	GPR polygons	G	72	7	
		Subtotals:	***94***	***13***	***0***
6	res lines	R	5		2
7	res polygons, high	R	7	1	
8	res polygons, low	R	12	2	
		Subtotals:	***24***	***3***	***2***
9	mag ditch, robust	D	6		2
10	mag ditch, subtle	D	5		1
11	GPR ditch	D	4		2
12	res ditch, high	D	4		1
13	res ditch, low	D	4		2
		Subtotals:	***23***	***0***	***8***
		TOTAL:	***279***	***30***	***11***
		AREA (m^2):		***240***	***120***

Abbreviations: mag = magnetic gradiometry, GPR = ground-penetrating radar, res = electrical resistance, M = mag stratum, G = GPR stratum, R = res stratum, D = "ditch" stratum, SB = 2 x 4 m strip blocks, BT = 1 x 8 – 1 x 20 m backhoe trench.

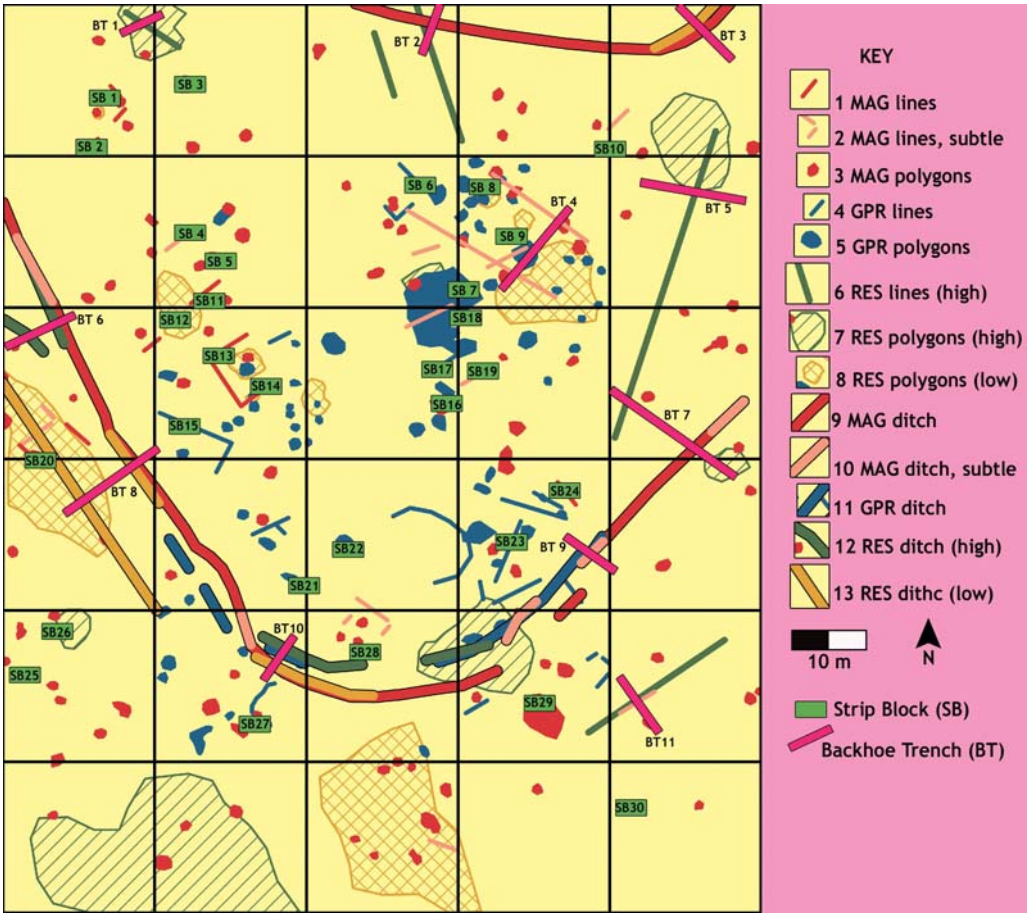


Figure 5.54. *Distribution of excavation units for Kasita Town overlaid on the vector interpretation. Linear anomalies and excavation unit symbols are exaggerated.*

5.15. VALIDATION PHASE FIELD METHODS

Michael L. Hargrave, Construction Engineers Research Laboratories, with contributions by Kenneth L. Kvamme, University of Arkansas

The need to excavate a representative sample of anomalies from numerous projected categories of archaeological types quickly placed challenging demands on the fieldwork budget. Funds allocated for this task at each site were roughly comparable to the upper end of the range in project costs for NRHP eligibility evaluations (ca. $25,000). Initial plans called for ground truthing investigations at only three of the sites (Kasita Town, Silver Bluff, and Pueblo Escondido). It was hoped that excavations that had been conducted previously at Army City (Kreisa and Walz 1996; Larson and Penney 1997) would provide an adequate sample for validating the remote sensing findings of the SERDP team. It later became apparent that the remote sensing results from Army City would be particularly useful in (1) developing and evaluating field protocols owing to very clear and recognizable anomaly forms (Section 4.6.1), and (2) that the extant excavations were neither a representative sample of the site's features nor large in number for an adequate evaluation. It was therefore decided to conduct ground truthing excavations at Army City as well. Yet, the need to investigate four sites using funds

initially set aside for three obviously reduced the amount of work that could be done at each, but it was apparent that this would be in the best interests of the overall project.

5.15.1. A Multi-Staged Strategy for Anomaly Validation

Prior to the SERDP project, CERL had found that a multi-staged approach to ground truthing could be highly effective, particularly in situations where only a relatively small percentage of anomalies could be investigated (Hargrave 2006). In a staged approach, ground truthing begins with the use of the *least* invasive and *least* expensive investigation techniques, with a goal of eliminating some anomalies from further consideration at each stage. Five stages are envisioned in this program (Ahler et al. 1999, 2003; Hargrave 2006).

1. Ground truthing is initiated with a careful surface inspection, and anomalies that could potentially be explained by the presence of such phenomena as vehicle ruts, tree roots, rodent burrows, military foxholes, etc., are eliminated.
2. At prehistoric sites, a metal detector is used to quickly examine robust magnetic or EM "point" or dipolar anomalies and eliminate those derived from metallic objects (which do not generally belong in the prehistoric period).
3. A one-inch diameter ("Oakfield") probe is used to extract soil cores from inside and outside each anomaly. Anomalies that exhibit soil characteristics *different* from those of their surroundings are retained in the sample whereas all others are eliminated.
4. The previous comparative approach is repeated using screened shovel test pits (STP), again from inside and outside of the anomaly to better evaluate whether real differences exist (in soil type, color, texture, and inclusions such as artifacts, charcoal, ash, etc.).
5. Those anomalies that survive the first four stages are investigated using 1 x 1 m excavation units for the ultimate evaluation.

The SERDP team initially planned to employ this staged approach for ground truthing. As sampling designs were being developed for the SERDP sites, it was realized that it would be difficult and unnecessary to use a staged approach. One goal of the staged approach is to reduce the number of anomalies to be investigated to a level that is consistent with the project budget. A second goal is to identify anomalies that most merit investigation, i.e., those judged most likely to be associated with features. The use of pre-selected random samples of anomalies, fixed in number, precluded the need to meet these goals. Moreover, maps that showed the pre-selected random sample for each site eliminated the need to plan for innumerable contingencies concerning which validation technique would need to be implemented before assessing each anomaly. This greatly facilitated the fieldwork requirements of each CRM consulting firm.

Nevertheless, the ideal situation would have been to investigate each of the selected anomalies using 1 x 1 m or larger hand-dug excavation units, but budgetary realities precluded this. Excavations at the sites included a mix of backhoe trenches usually three meters long by .75 meters in width, 1 x 0.5 m "hand excavation" units, and "shovel test pits" (STP) measuring about 35 x 35 cm (Figure 5.55a-c). One-inch soil augers or corers were also employed to augment knowledge gained in excavations, for

example, to ascertain the lateral and vertical extents of discovered archaeological features (Figure 5.55d). This panoply of field evaluation methods thus provided an opportunity to evaluate the effectiveness of alternative excavation techniques (in terms of both cost and information return).

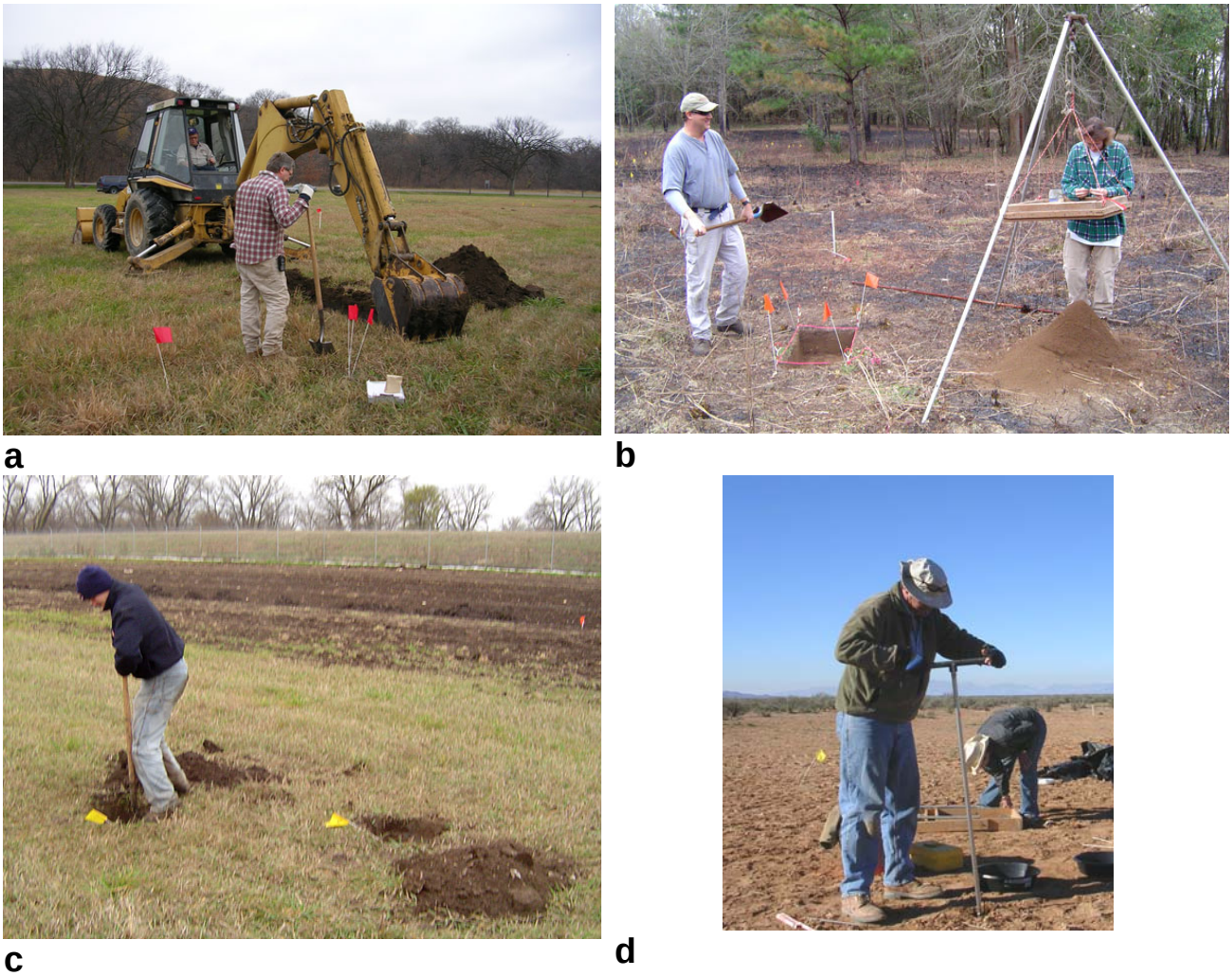
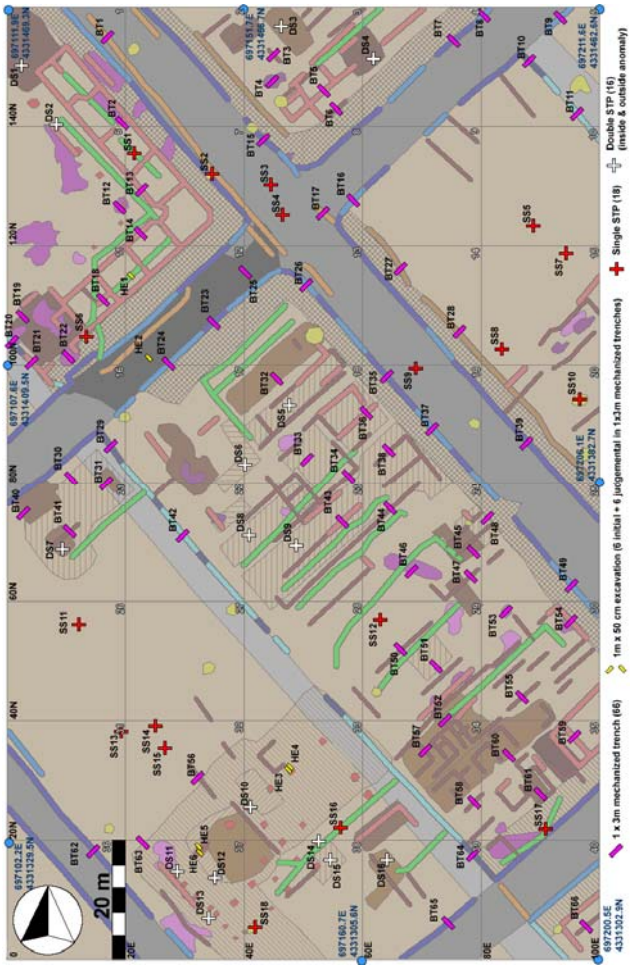


Figure 5.55. *Excavation methods used for field validation in the SERDP Project. a) Backhoe trench (typically 3 x .75 m), b) hand excavation (typically 1 x .5 m), c) STP (typically .35 x .35 m), and d) one-inch Oakfield soil core.*

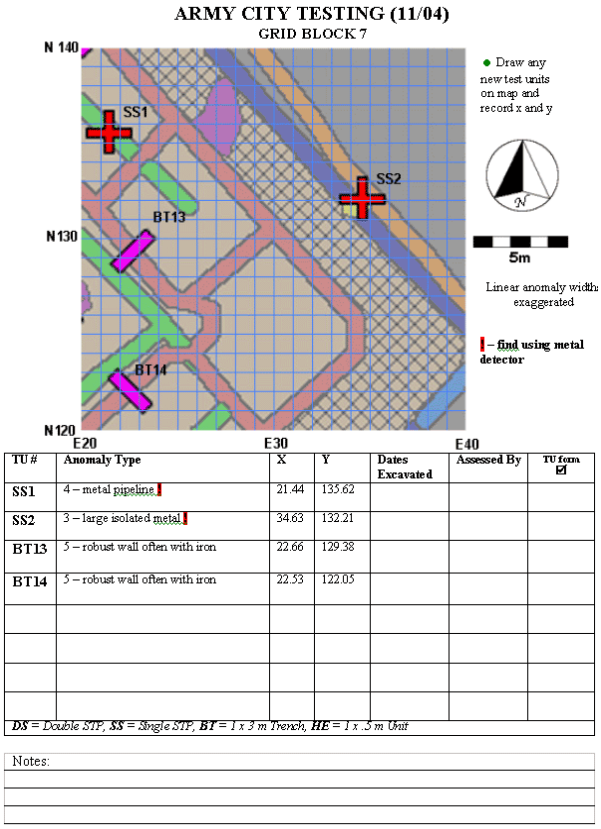
5.15.2. The SERDP Project's Field Validation Strategy

Ground truthing at each of the four sites employed a standardized protocol, although details were modified in response to local conditions. The site sampling designs (Section 5.14) assigned anomalies to a number of categories that reflected archaeological or descriptive types of subsurface features, and a random sample of anomalies was drawn from each. Decisions were made about the type of excavation unit that should be used to investigate each anomaly and the unit's location relative to the anomaly. These decisions had to be made within the context of budget limitations (approximately $18,750 was available for each site). It was also desirable to use minimally invasive techniques such as

soil probes, augers, and STPs to minimize costs as well as impacts to the sites. Yet, it was feared that such techniques would often not provide sufficient information. It therefore was decided to focus on mechanized trenching as a reasonably economical way to expose the subsurface sufficiently to facilitate recognition of archaeological features, their composition, and dimensions. As indicated in Figure 5.56a, each excavation unit was numbered, their locations were added to the interpreted anomaly maps (see Section 5.14), and each anomaly category and excavation type was color-coded. Moreover, a one-page form for each 20 x 20 m study block showed a close-up map of anomalies included in the sample, the loci of excavation units, and listed grid coordinates for at least one corner of each of the planned excavations (Figure 5.56b).



a



TU #	Anomaly Type	X	Y	Dates Excavated	Assessed By	TU form ☑
SS1	4 – metal pipeline	21.44	135.62			
SS2	3 – large isolated metal	34.63	132.21			
BT13	5 – robust wall often with iron	22.66	129.38			
BT14	5 – robust wall often with iron	22.53	122.03			

DS = Double STP, SS = Single STP, BT = 1 x 3 m Trench, HE = 1 x .5 m Unit

Notes:

b

Figure 5.56. *Forms used in field evaluations at Army City. a) Map showing distribution of all interpreted anomalies and excavation units for field evaluations. b) Field form for a 20 m study block showing excavation units in that block.*

The actual fieldwork at each site was conducted by a CRM consulting firm with expertise in the region in conjunction with the SERDP project team. Assistance from local research consultants was essential for two reasons. First, each of the project's sites had already been found to be eligible for nomination to the National Register of Historic

Places. As such, each site represented an important source of scientific and cultural information. The only way to justify the adverse impacts that would result from archaeological excavation was to prepare professional-quality reports of the investigations for distribution to the host installations' CRM program, the relevant State Historic Preservation Office (SHPO), and other professional archaeologists and interested members of the public. Standard professional practice also required the SERDP project to ensure that the recovered artifacts and associated records were prepared for long-term curation in conformance with federal standards. The SERDP researchers lacked the time, regional, and topical expertise to personally fulfill all of these responsibilities.

The SERDP team also recognized the importance of independent, objective evaluations of their interpretations of the remote sensing data. The categories that archaeologists use to describe and interpret cultural features often allow considerable room for subjective assessments, and this projected wanted to capture and recognize that variation. What one investigator interprets as a house wall may, for example, be described as a linear soil stain by another researcher. Additionally, the SERDP team members were more familiar than most archaeologists with how variations in soil characteristics, rocks, bioturbations, and other phenomena might be manifested in the remote sensing results. It was important to demonstrate that the SERDP team's interpretations and field observations could, at least in general terms, be corroborated by colleagues with expertise in the local archaeology, but with little or no previous experience in remote sensing.

Qualified firms were identified based on recommendations from the installation Cultural Resource Managers, who were familiar with the capabilities and work history of CRM firms in their area. The SERDP team prepared detailed Statements of Work (SOWs) for each ground truthing project and contracts were awarded through ERDC CERL. The SOWs were designed to ensure that the contractors understood the SERDP project's goals and data needs, and use the excavation techniques selected by the project team. However, the SOWs were also written to allow sufficient latitude to respond to actual field conditions. The contractors were responsible for the field investigations, a thorough analysis and interpretation of the artifacts and contextual data, the proper preparation of the artifacts for long-term curation, and the completion of a detailed written report.

5.15.3. Field Preparation and Fieldwork

The SERDP team arrived at each site a day or two before the contractors to mark the exact locations of all excavation units using non-metallic PVC pin-flags (Figure 5.57a). This was accomplished through use of high-accuracy global positioning systems to relocate buried datums from the geophysical survey work of 2002 followed by use of transits and tape measures to locate anomalies relative to 20 m survey block corners. Location accuracy proved to be high and probably within 10 cm in all cases. Pin flags were labeled with unit designations (e.g., MT1 for "Mechanized Trench 1"). The consulting firm's field director was provided with a complete set of the anomaly and excavation unit maps described above (Figure 5.56), and requested to use the unit and anomaly designations that had been assigned by the SERDP team in all field records and throughout the written report.

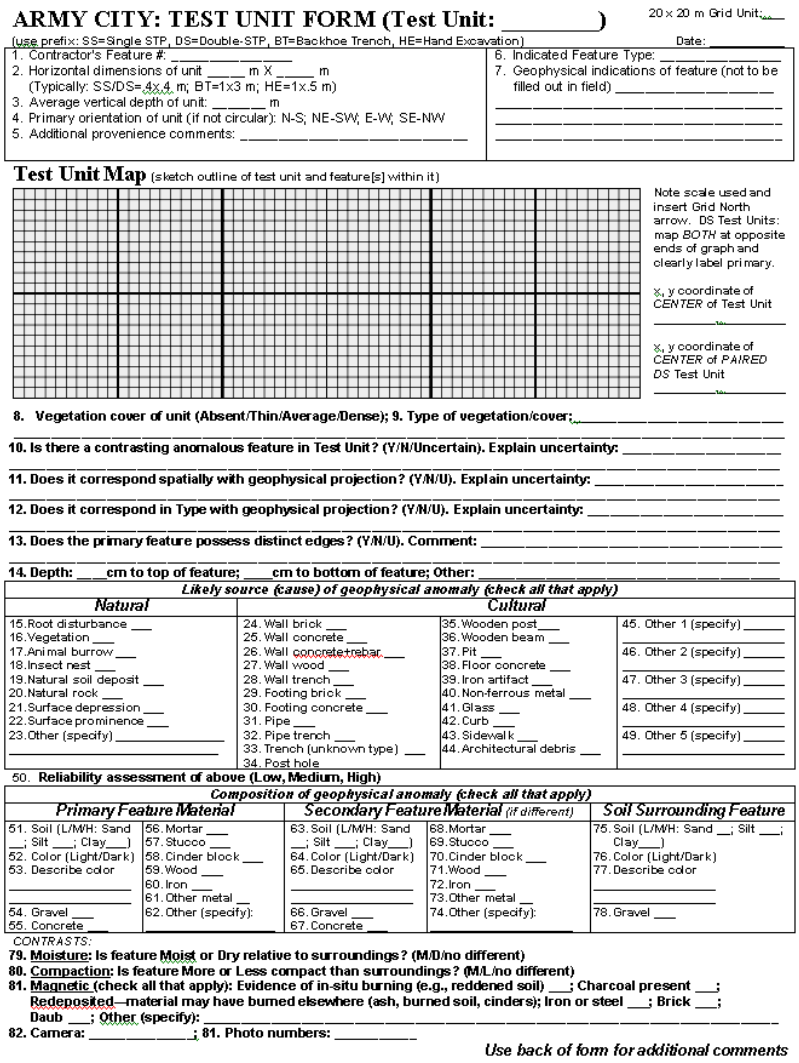
a **b**

Figure 5.57. *Validation phase fieldwork. a) Placement of PVC pin flags at Army City to precisely locate excavation units prior to excavations. b) Archaeological team (foreground) and SERDP team (background) making independent evaluations of excavation findings at Pueblo Escondido.*

The consulting firm's fieldwork was scheduled to be completed in approximately five working days and this was achieved at all sites. In the field the SERDP project team and the contractor's crew interacted closely, but had distinct responsibilities. The contractors were responsible for the excavation, documentation, and back-filling of each unit, and for ensuring that all artifacts and project records were recovered with proper provenience. The maps provided by the SERDP team indicated the instrument that primarily revealed each anomaly and, in many cases, the inferred type of archaeological feature expected to be present. The contractors were the first to visit each unit and were responsible for recording any evidence for such features, as well as for any other features or archaeologically important deposits. If no features were present, they were urged to describe any evidence, even if very subtle, for unusual conditions (bioturbations, localized dips in soil strata, etc.) that might account for the anomaly. Each of the contractors made a good faith effort to do this, but given the many tasks they had to complete, time and budgetary limitations did not permit sufficient time for careful scrutiny of units that were, by conventional standards, devoid of cultural features. Throughout this process the SERDP team was careful not to bias the archaeologist's interpretations of subsurface features, stratigraphic content, or anomaly sources in order that their independent evaluations could be derived.

As noted previously, the SERDP team independently documented each excavation unit (Figure 5.57b). They took detailed notes, made sketch maps and photographs of each unit, recorded observations on standardized forms (Figure 5.58), and focused specifically on identifying possible sources for the anomalies, particularly in situations where no archaeological feature (as the term is typically used) appeared to be present. The SERDP team generally was able to spend much more time considering this task than the archaeological teams, even resorting to cutting back or re-facing trench walls and floors for a clearer view of soil and other variations or using soil augers to check depths and extents of visible features. In some cases, a hand-held magnetic susceptibility meter (the KT-9 Kappameter) was used to better understand soil strata and

possible features (except at Army City). At Kasita Town, a hand-held probe array was also used with a Geoscan Research RM15 electrical resistivity meter to map the wall of one trench. These efforts provided data to better understand how subtle, localized variations in soil texture, depth of soil strata, and the presence of roots, rocks, or other phenomena were manifest in the data recorded by the various sensors.



ARMY CITY: TEST UNIT FORM (Test Unit: _____) 20 x 20 m Grid Unit:___

(use prefix: SS=Single STP, DS=Double-STP, BT=Backhoe Trench, HE=Hand Excavation) Date: _____

1. Contractor's Feature #: _____
2. Horizontal dimensions of unit ____ m X ____ m (Typically: SS/DS=.4x.4 m; BT=1x3 m; HE=1x.5 m)
3. Average vertical depth of unit: _____ m
4. Primary orientation of unit (if not circular): N-S; NE-SW; E-W; SE-NW
5. Additional provenience comments: _____
6. Indicated Feature Type: _____
7. Geophysical indications of feature (not to be filled out in field) _____

Test Unit Map (sketch outline of test unit and feature[s] within it)

Note scale used and insert Grid North arrow. DS Test Units: map *BOTH* at opposite ends of graph and clearly label primary.

x, y coordinate of *CENTER* of Test Unit _____

x, y coordinate of *CENTER* of *PAIRED DS* Test Unit _____

8. Vegetation cover of unit (Absent/Thin/Average/Dense); 9. Type of vegetation/cover: _____

10. Is there a contrasting anomalous feature in Test Unit? (Y/N/Uncertain). Explain uncertainty: _____

11. Does it correspond spatially with geophysical projection? (Y/N/U). Explain uncertainty: _____

12. Does it correspond in Type with geophysical projection? (Y/N/U). Explain uncertainty: _____

13. Does the primary feature possess distinct edges? (Y/N/U). Comment: _____

14. Depth: ___cm to top of feature; ___cm to bottom of feature; Other: _____

Likely source (cause) of geophysical anomaly (check all that apply)

Natural	*Cultural*		
15. Root disturbance ___	24. Wall brick ___	35. Wooden post___	45. Other 1 (specify) _____
16. Vegetation ___	25. Wall concrete ___	36. Wooden beam ___	
17. Animal burrow ___	26. Wall concrete+rebar ___	37. Pit ___	46. Other 2 (specify) _____
18. Insect nest ___	27. Wall wood ___	38. Floor concrete ___	
19. Natural soil deposit ___	28. Wall trench ___	39. Iron artifact ___	47. Other 3 (specify) _____
20. Natural rock ___	29. Footing brick ___	40. Non-ferrous metal ___	
21. Surface depression ___	30. Footing concrete ___	41. Glass ___	48. Other 4 (specify) _____
22. Surface prominence ___	31. Pipe ___	42. Curb ___	
23. Other (specify) _____	32. Pipe trench ___	43. Sidewalk ___	49. Other 5 (specify) _____
	33. Trench (unknown type) ___	44. Architectural debris ___	
	34. Post hole ___		

50. Reliability assessment of above (Low, Medium, High)

Composition of geophysical anomaly (check all that apply)

Primary Feature Material		*Secondary Feature Material (if different)*		*Soil Surrounding Feature*
51. Soil (L/M/H: Sand __; Silt ___; Clay___)	56. Mortar ___	63. Soil (L/M/H: Sand __; Silt ___; Clay___)	68. Mortar ___	75. Soil (L/M/H: Sand __; Silt ___; Clay___)
52. Color (Light/Dark)	57. Stucco ___	64. Color (Light/Dark)	69. Stucco ___	76. Color (Light/Dark)
53. Describe color _____	58. Cinder block ___	65. Describe color _____	70. Cinder block ___	77. Describe color _____
	59. Wood ___		71. Wood ___	
	60. Iron ___		72. Iron ___	
	61. Other metal ___		73. Other metal ___	
54. Gravel ___	62. Other (specify): _____	66. Gravel ___	74. Other (specify): _____	78. Gravel ___
55. Concrete ___		67. Concrete ___		

CONTRASTS:

79. Moisture: Is feature Moist or Dry relative to surroundings? (M/D/no different)

80. Compaction: Is feature More or Less compact than surroundings? (M/L/no different)

81. Magnetic (check all that apply): Evidence of in-situ burning (e.g., reddened soil) ___; Charcoal present ___; Redeposited—material may have burned elsewhere (ash, burned soil, cinders); Iron or steel ___; Brick ___; Daub ___; Other (specify): _____

82. Camera: _____; 81. Photo numbers: _____

Use back of form for additional comments

Figure 5.58. *Excavation unit form used by SERDP team to evaluate archaeological sources of remote sensing anomalies at Army City.*

In general, accuracy statistics of the SERDP team assessments tend to be highest, but they had the advantage of knowing how various remote sensing devices respond to subsurface conditions and therefore better knew what to look for in seeking explanations. The archaeological teams, with less available time, a focus primarily on archaeological evidence, and lacking in-depth understanding of remote sensing, were less sensitive to the many nuances of the record and their remote sensing impacts. Nevertheless, it is emphasized that the bulk of remote sensing assessments in future archaeological projects are likely to be made by archaeologists. Until training in archaeological remote sensing becomes commonplace in university archaeology programs, the accuracy assessments made by the archaeologists in this project will remain most representative of remote

sensing assessments. Results of the SERDP team's evaluation of the excavations are presented with the individual contractor's findings to contrast perceptions of performance.

Results of the field investigations are summarized below for each site, based on the contractor's draft reports. The need to work with drafts was anticipated given the project schedule, and the SOWs specified that, in terms of content, the draft reports should be essentially complete. Army City was the first site to be investigated and in retrospect, that work went very smoothly. By comparison, each of the other three sites presented certain challenges and frustrations. The archaeological deposits at Pueblo Escondido were deeper than anticipated and in many cases the features were clearly present but difficult to delimit. Field conditions were challenging, including two days of intermittent heavy rain and a brief hail storm. Evaluating the results of ground truthing at Silver Bluff was complicated by the occurrence of many linear anomalies and an archaeological record dominated by post holes, middens, and amorphous stains. At Kasita Town, excavations revealed the site had been graded, resulting in the loss of near-surface deposits that truncated or eliminated many archaeological features in some areas and the occurrence of a thin layer of spoil in others. The consultant's field director apparently did not fully understand his responsibility to independently evaluate the degree to which features identified in the excavations could explain the occurrence of the remote sensing anomalies. As a consequence, it is necessary to rely primarily on the SERDP team's assessment of the accuracy of the remote sensing findings.

5.16. ARCHAEOLOGICAL INVESTIGATION OF REMOTE SENSING PREDICTIONS

Previous sections of this study have described remote sensing method and theory, computer and field methods, site histories and remote sensing results, a host of data fusion techniques, interpretations, and sampling designs for validating findings learned through these efforts. This penultimate section confronts the issue of validating or "ground truthing" results projected from the remote sensing and data fusion programs, utilizing the sampling designs of Section 5.14 and field methods described in Section 5.15. The focus rests very much in the techniques of archaeological excavation where two questions receive primary attention:

1. Is there observable physical evidence in the ground at the locus of an anomaly that is the likely source for that anomaly?
2. Is the physical evidence in the ground consistent with the interpretation of that anomaly?

The first question seeks to establish that an archaeological feature truly exists—that an anomaly does not represent one of the many "false positives" characteristic of remote sensing. The second attempts to ascertain whether the classification of an anomaly to a particular type is accurate—is an anomaly indicated to represent a wall truly associated with a wall or some other class of archaeological feature, such as a narrow ditch? Both questions relate to accuracy. On the one hand, remote sensing's ability to merely indicate an archaeological feature of unknown type is useful and important; on the

other, it is worthwhile to learn whether it might be able to indicate specific types of archaeological circumstances hidden beneath the soil. Cases where remote sensing projections are wrong contribute to a learning process that improves subsequent interpretations. They are used in Section 5.17 to offer a revision of interpretations for the project sites.

The following sections attempt to examine remote sensing accuracy, but under the limitations of archaeological excavation methods. As discussed in Section 5.13, field archaeology is an inexact science where uncertainty is frequently the norm. While excavating and exposing easy-to-recognize brick walls or hearths as a means of validating corresponding anomalies may seem straightforward, too often archaeologists encounter only stains in the soil and great uncertainty about what they might represent. Whether a particular stain might represent an indicated floor, a burned area, or a midden, may be open to question. In rare instances sources of anomalies may not be visible to excavators, who are limited only to senses of sight and touch compared to remote sensing devices. Clear examples of significant archaeological features were illustrated in Section 5.13 that left no macroscopically visible evidence in the excavations. Taken together, these issues contribute to sources of error in excavation-based validation that must always be considered. Ultimately, as the art and science progress, various laboratory testing methods and geoarchaeological techniques may also be incorporated into evaluation methodologies for assessing remote sensing accuracy.

5.16.1. Field Investigations at Army City

Michael L. Hargrave, Construction Engineers Research Laboratories (CERL) & Kenneth L. Kvamme, University of Arkansas

Excavations at Army City were conducted by the University of Illinois's Public Service Archaeology Program (PSAP), under the direction of Dr. Paul Kresja. PSAP was selected for this work because of their history of high quality, cost-effective work at Fort Riley. Dr. Kresja directed field investigations in the northeastern portion of Army City in 1996 (see Section 3.3.4.1), and was thus very familiar with the site's soils and the nature of its archaeological deposits (Kreisa and Walz 1996; note that in recent years Dr. Kresja has chosen to use the traditional spelling of his family name). The excavations were conducted in November, 2004, by Dr. Kresja and a crew of three. Three members of the SERDP team (Kvamme, Ernenwein, Hargrave) and two student research assistants from the University of Arkansas were also present. At Army City the SERDP team shared responsibility with the contractor's field director in monitoring excavations by the backhoe (Figure 5.59). Once all of the planned trenches had been excavated, the SERDP team and PSAP crew independently documented each of the units.

Figure 5.59. *View of project excavations at the site of Army City, looking south from nearby bluffs in November 2004.*

The remote sensing study area at Army City consists of forty blocks 20 x 20 m in size that covered a total area of 16,000 m^2 (1.6 ha or 3.95 acres; see Section 4.6.1). Following the sampling plan outlined in Table 5.20 (Section 5.14.1), excavations included 66 trenches excavated by a backhoe equipped with a toothless, .75 m-wide bucket. Plans called for each trench to be approximately 3 meters long by one meter wide but, in reality, most trenches were slightly longer and only .75 m wide (the width of the bucket). The backhoe excavated in a series of shallow scraping cuts, removing 5-10 cm at a time, allowing field personnel to carefully monitor the appearance of changes in deposits, archaeological features and artifacts, and that it did not inadvertently cut through a feature or significant deposit. Excavation ceased when a feature was encountered or, in the absence of features, when subsoil clearly below archaeological deposits was reached (Kresja 2005:23).

The entire sampling plan of 106 excavation units was fully carried out at Army City (Table 5.20). Fifty shovel test pits, each measuring at least 35 x 35 cm in plan, were excavated in 10 cm levels. All hand-excavated soil was screened through 6.35 mm mesh. Like the trenches, STPs were excavated until either an archaeological feature or subsoil was encountered. Finally, four excavation units, measuring 1 x .5 m or 1 x 1 m, were hand excavated using the same protocol as the STPs. PSAP documented all test units, trenches, and probes by scaled maps, digital photographs, and standardized excavation forms (Kresja 2005:24).

As was suggested by earlier excavations at Army City (Kreisa and Walz 1996), the site holds a plethora of archaeological remains of identifiable type. These include concrete floors and walls, burned wooden walls, masonry footings, a fireplace base, cement sidewalks, ceramic and steel pipes, and even hints of former street gutters (Figure 4.37). Collectively, they offer excellent opportunity to test the remote sensing predictions of Section 5.12.4 through the process of comparing their correspondences with actual archaeological circumstances (although the latter cannot faithfully be determined in some cases, as discussed in Section 5.13.1). The following sections make this assessment from two perspectives, one from the SERDP team and one from the archaeological team, each with its own set of expertise and biases.

5.16.1.1. Evaluation of predictions by archaeological team
Michael L. Hargrave, CERL

Kresja (2005:185) used data from the excavations and observations of his field team to evaluate the accuracy of remote sensing predictions in three ways: 1) the overall percentage of correct predictions (i.e., an archaeological feature was documented where predicted) for all anomaly and excavation unit types; 2) the percentage of correct predictions by type of excavation unit (STP, hand excavation, or backhoe trench); and 3) the percentage of correct predictions for each anomaly type. In the following, percentages were calculated using 96 of the total of 120 excavation units. Twenty-four STPs, intentionally excavated at locations where no anomalies were present and no features were anticipated, are excluded in the following.

Overall, 69.8 percent (67 of 96) of the investigated anomalies were found to be associated with the type of feature predicted to be present (Table 5.24). An additional 16.7% (16 of 96) of the anomalies were also reported to be associated with features, but not with the type of feature that had been predicted. Note, however, that this category does not imply that one of the other feature types defined by SERDP was present. In almost all cases, the "other feature" was described as a soil or fill zone or a concentration of fragmented concrete, coal, or other material. In sum, 86.5% of the anomalies were associated with an archaeological feature. No feature was identified in the remaining 13.5% (n=13) of the units (Kresja 2005:185).

Table 5.24. *Overall percentage of correct predictions for all excavation units (adapted from Kresja 2005 Table 10). "Other feature type" generally refers to a soil or fill zone, not a formal feature.*

Feature Type	Predicted Feature Type Present		Other Feature Type Present		No Feature Present		Total
	N	%	N	%	N	%	N
Ruts or grooves	1	50.0%	1	50.0%	0	0.0%	2
Walls	16	66.7%	2	8.3%	6	25.0%	24
Debris or rubble	4	80.0%	0	0.0%	1	20.0%	5
Gutters or curbs	13	65.0%	6	30.0%	1	5.0%	20
Floors or surfaces	17	77.3%	1	4.5%	4	18.2%	22
Sidewalk edges	2	100.0%	0	0.0%	0	0.0%	2
Streets	1	20.0%	3	60.0%	1	20.0%	5
Footers or chimney bases	7	87.5%	1	12.5%	0	0.0%	8
Metal	6	75.0%	2	25.0%	0	0.0%	8
Total	67	69.8%	16	16.7%	13	13.5%	96

It seemed likely that the type of unit excavated would have an effect on the identification of archaeological features. For example, the backhoe trenches typically exposed a horizontal area of approximately 2.25 m^2, and the trench walls provided substantial additional opportunities to identify features that might not be discernable in the trench floor (Figure 5.60). In contrast, the shovel tests each exposed a horizontal area of only about .123 m^2 and the walls provided far less opportunity to detect a feature than did the walls of a backhoe trench. In addition to the difference in area exposed, the shovel tests offered much less opportunity to closely examine subtle differences in soil color and texture due to limited lighting and access. In the case of relatively small anomalies, it was clearly essential that the shovel tests were properly located in the field. Note that in the following discussion, Kresja combined the four larger hand excavation units (of 1 x .5 m,

which each exposed a horizontal area of .5 m^2) with the 26 shovel tests that were targeted on anomalies (Kresja 2005:186).

Figure 5.60. *Trench 12 at Army City was excavated to investigate a robust concentration of debris or rubble. A large concrete block was documented in the east (right site) wall, November 2004.*

Surprisingly, there appears to be no difference between backhoe trenches and the shovel tests combined with hand-excavation units in terms of the percentage of accurate predictions (Tables 5.25 and 5.26). Considering all categories together, 70% of the units encountered features of the type predicted to be present. Among the shovel tests and hand-excavation units, 20% produced a feature that was not the type predicted. In the remaining 10%, no feature was identified. For the backhoe trenches, 15% of the units contained a feature that was not the type predicted, and the remaining 15% produced no indications of any feature (Kresja 2005:187).

Table 5.25. *Percentage rate of correct predictions for shovel tests and hand-excavation units (combined) (adapted from Kresja 2005:Table 11)*

Feature Type	Predicted Feature Type Present		Other Feature Type Present		No Feature Present		Total
	N	%	N	%	N	%	N
Floors or surfaces	7	70.0%	1	10.0%	2	20.0%	10
Footers or chimney bases	7	87.5%	1	12.5%	0	0.0%	8
Metal	6	75.0%	2	25.0%	0	0.0%	8
Street surfaces	1	25.0%	2	50.0%	1	25.0%	4
Total	21	70.0%	6	20.0%	3	10.0%	30

Table 5.26. *Percentage rates of correct predictions for backhoe trenches (adapted from Kresja 2005:Table 12).*

Feature Type	Predicted Feature Type Present		Other Feature Type Present		No Feature Present		Total
	N	%	N	%	N	%	N
Ruts or grooves	1	50.0%	1	50.0%	0	0.0%	2
Walls	16	66.7%	2	8.3%	6	25.0%	24
Debris or rubble	4	80.0%	0	0.0%	1	20.0%	5
Gutters or curbs	13	65.0%	6	30.0%	1	5.0%	20
Floors or surfaces	10	83.3%	0	0.0%	2	16.7%	12
Sidewalk edges	2	100.0%	0	0.0%	0	0.0%	2
Streets	0	0.0%	1	100.0%	0	0.0%	1
Total	46	69.7%	10	15.2%	10	15.2%	66

The apparent lack of a difference between excavation unit types in terms of the rate of accurate predictions about features is probably due to the non-random manner in which decisions about which type of unit to use were made. For example, the hand-excavation units were used to verify anomalies that were confidently predicted to be associated with features such as building footers (Figure 5.61). Backhoe trenches were generally selected to investigate anomalies predicted to be associated with features that might be difficult to identify, such as ruts, grooves, and gutters. Shovel test pits would almost certainly have been less useful than trenches in efforts to verify the presence of such features. A reliable comparison of the usefulness of different types of units would have required a randomized relationship between unit type and anomaly/feature type.

Figure 5.61. *Concrete footer identified in a test unit at Army City, November 2004.*

Finally, one would expect the anomaly types to differ in terms of the percentage of correct predictions. Nine types of anomalies were identified in the geophysical data from Army City: road ruts and grooves, structure walls, debris and rubble fields or concentrations, street and alley gutters or curbs, floors or surfaces, sidewalk edges, street surfaces, building footers, chimney bases, and metal objects (including both pipes and isolated pieces of metal). In this evaluation, excavation units that produced evidence for a feature other than the type predicted to be present are reported only as "negative findings," as are those units that produced no feature (Table 5.27) (Kresja 2005:187).

Table 5.27. *Incidence of positive and negative results in efforts to verify the presence of features associated with each anomaly type.*

	Ruts, grooves	Gutters, curbs	Sidewalk edges	Street surface	Walls	Floors, surfaces	Debris, rubble	Metal	Footer, Chimney
Positive	1	13	2	1	16	17	4	6	7
Negative	1	7	0	4	8	5	1	2	1

Notes: Negative cases include those where an "other feature type" was present (generally a soil or fill zone) and cases where nothing to explain the anomaly was present

Kresja (2005:187) notes that two anomaly categories were particularly difficult to identify correctly in the geophysical data or in the field. These were classified as street surfaces and ruts or grooves. Two other anomaly categories—gutters or curbs, and walls—were also correctly predicted at a rate of less than the 70% that was achieved for all anomaly types combined (Table 5.24). Three of these categories (street surfaces, ruts or grooves, gutters and curbs) are linear and associated with streets and alleys. It is disappointing but perhaps not too difficult to understand why these linear categories were not verified as consistently as the other anomalies. It may be that linear anomalies are readily detected in a geophysical map because the human eye and mind tend to seek out coherent shapes and patterns. Additionally, the linear anomalies in question tend to co-occur in a consistent manner, contributing to the ease with which they are discerned in a geophysical image. In other words, street gutters and curbs co-occur with streets, they are not randomly distributed. These subtle anomalies are, however, very difficult to detect in the field within the confines of a relatively small trench. Army City's streets were unpaved, and anomalies such as gutters and ruts were, in many cases, probably filled with soil from the surrounding streets, resulting in very little visual or textural contrast.

Kresja (2005:188) reports that some anomaly types (sidewalk edges, debris or rubble concentrations, and building footers or chimney bases) were relatively easy to identify in the field. The latter two categories are comprised of relatively large pieces of concrete and or brick and are difficult to miss even within the confines of a small excavation unit. Archaeo-geophysicists would also assume that metal objects should be easy to identify both in the geophysical data and in the field. Only 75% of such anomalies were verified, but Kresja notes that this percentage would have been higher if ceramic tiles were included with metal pipes (Kresja 2005:188). Iron pipes and fired clay tiles are both characterized by relatively high amplitude magnetic values, making them potentially difficult to differentiate in a magnetic map.

Finally, Kresja notes that only two anomaly categories, floors or surfaces and street surfaces, were investigated using both backhoe trenches and shovel tests or test units (combined). Only one street surface was investigated using a backhoe trench, so that sample is too small to be meaningful. However, 83% of the floor or surface anomalies that were investigated using backhoe trenches were verified, whereas only 70% of those investigated using shovel test pits were confirmed. This appears to support the common-sense expectation that using larger excavation units will increase the likelihood of verifying the presence of aerially extensive anomalies such as those associated with floors or surfaces (Kresja 2005:188).

The Fort Riley CRM office was initially reluctant to permit the SERDP project to use a backhoe to ground truth numerous anomalies at Army City, and this was

understandable given the site's research value. In retrospect, however, a combination of geophysical survey and carefully targeted mechanized ground truthing appears to have been an ideal approach for investigating the site. Army City's archaeological remains were found to be dominated by massive deposits of construction debris and relatively small quantities of artifacts related to activities such as food consumption, commercial and recreational activities. Achieving a satisfactory understanding of the site's spatial character based solely on hand excavation (without geophysics) would have been enormously expensive. A combination of geophysical survey and carefully targeted mechanized excavation is likely to be a responsible, cost effective approach for evaluating the nature and integrity of the World War 1 and World War 2 era complexes of barracks and related buildings that exist at many U.S. military installations.

5.16.1.2. Evaluation of predictions by SERDP team
Kenneth L. Kvamme, University of Arkansas

This section reviews the nature of the archaeological evidence for each defined anomaly type of Table 5.16. It does so by presenting the primary geophysical evidence for a sample of each type and illustrating archaeological sources for the anomalies indicated. The relative ease or difficulty of locating archaeological sources for each type is also discussed. Finally, accuracy of anomaly identification is presented based on the SERDP team's field evaluations and data. Accuracy is evaluated by anomaly type, excavation type, and other means. Some of the following material differs from the foregoing assessments by the archaeological team. They lumped isolated metal artifacts and pipes into a single category, for example. Moreover, two excavations intended to sample the "background" (loci without anomalies) were placed by random selection within unmodified areas of the town's streets. The archaeological team erroneously thought they were placed to locate a former street surface. Other similar examples exist.

The anomaly types "chimney base" and "footings" were robustly indicated by the electrical resistivity survey (Figure 5.62a,c), with subtle indications in some of the other surveys. The interpretation of the chimney base came from a somewhat larger and rectangular anomaly that was linked with an historic photo that actually illustrates the chimney (Figure 3.2; structure in right background). Two adjacent hand excavation units, HE5 and HE6, were employed to expose and validate this unique archaeological feature (Figure 5.62d). The interpretation of building footings was based on the size, shape, and robust contrast of multiple resistivity anomalies, but primarily on their spatial distributions in systematic rows and columns (Figure 5.62a,c). Several were exposed in the field evaluation and were found to be composed of concrete, brick, or stone. Small STP were found to be an effective means of evaluation for this easily recognized feature type (e.g., Figure 5.62f), but a rather large one that once supported the Hippodrome's superstructure was revealed in the side of a backhoe trench that also shows an associated builder's trench and different soils about its perimeter (Figure 5.62a,b). A pair of adjacent hand excavation units, HE3 and HE4, was employed to fully expose one complete footing for documentation purposes (Figure 5.62e). The ability to physically locate these small archaeological features and place archaeological excavation units precisely over them emphasizes the high level of care and accuracy undertaken during the data collection phase of the project. These classes of anomalies represent "high confidence"

classes generally proven to represent correct targets in excavation (see below), although one suspected footing turned out to by a collection of thin slabs of stucco.

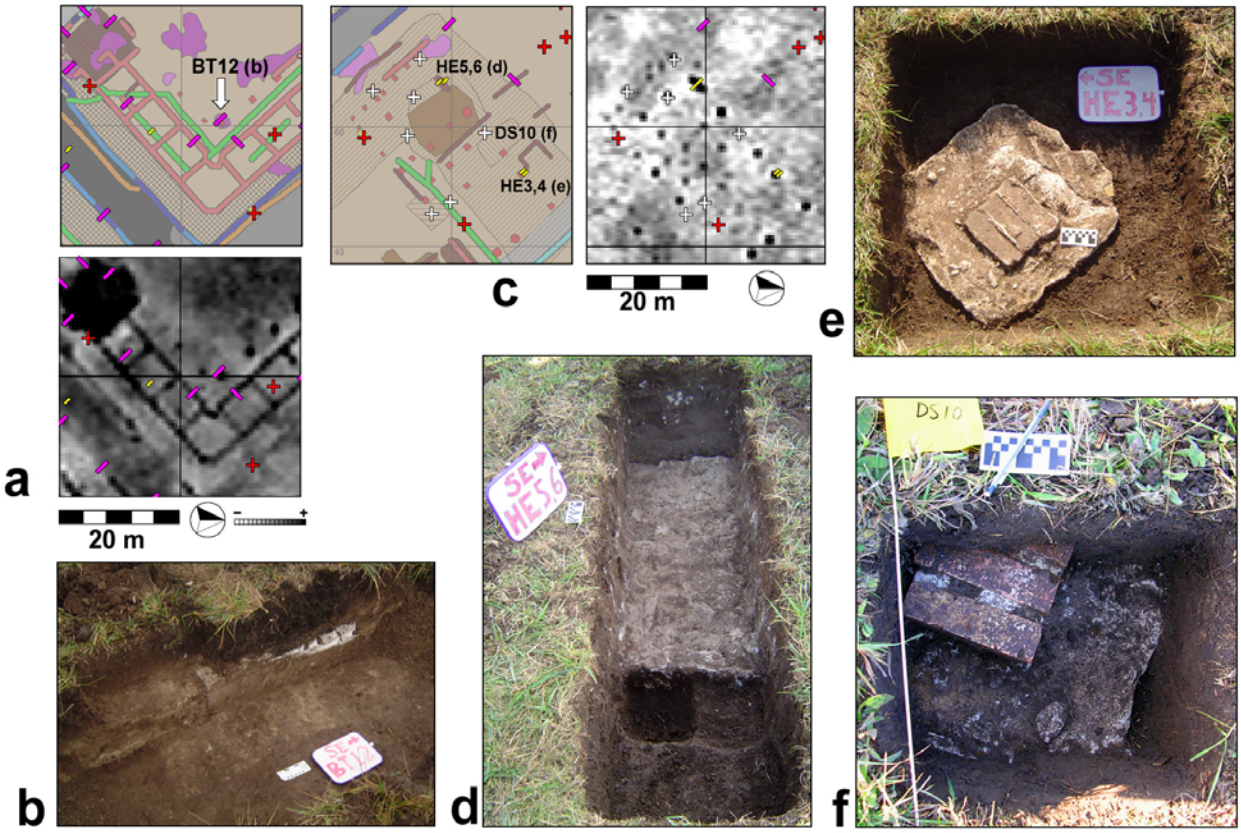


Figure 5.62. *Samples of archaeological sources for anomalies classified as "chimney base" and "footings." a) Vector interpretations and resistivity image showing locus of anomaly interpreted as a large footing in the Hippodrome and backhoe trench BT12. b) Backhoe trench BT12 showing footing in profile and associated builder's trench. c) Vector interpretations and resistivity image showing loci of anomalies interpreted as a chimney base (HE5,6) and footings with excavation units indicated. d) Hand excavations HE5,6 exposing a chimney base. e) Hand excavations HE3,4 revealing a complete footing of concrete with brick impressions. f) STP DS10 showing footing with bricks. KEY: INTERP=interpretations; RES=resistivity; BT=backhoe trench (purple); DS=double STP (white); HE=hand excavation (yellow).*

An historic 20th century site such as Army City contains numerous metal artifacts, each producing anomalies, and generally robust ones (Figure 5.63a) in several of the geophysical data sets (Section 4.6.1). It was felt that small metal-generated "point" anomalies were a legitimate class to test, but excessive focus was not given to them. Rather, most anomalies of this class were screened out leaving only the most robust or largest ones. This left two distinct "metallic" anomaly types for investigation: "large isolated metal artifacts" and "pipes." Several of the former were tested employing simple STP because of ease of location (assisted by a metal detector). One, SS2, turned out to be a Kansas license plate (with missing date due to corrosion) lying in a street gutter (Figure 5.63b). A second, SS18, was revealed to be a steel lightening rod placed vertically in the ground (Figure 5.63c).

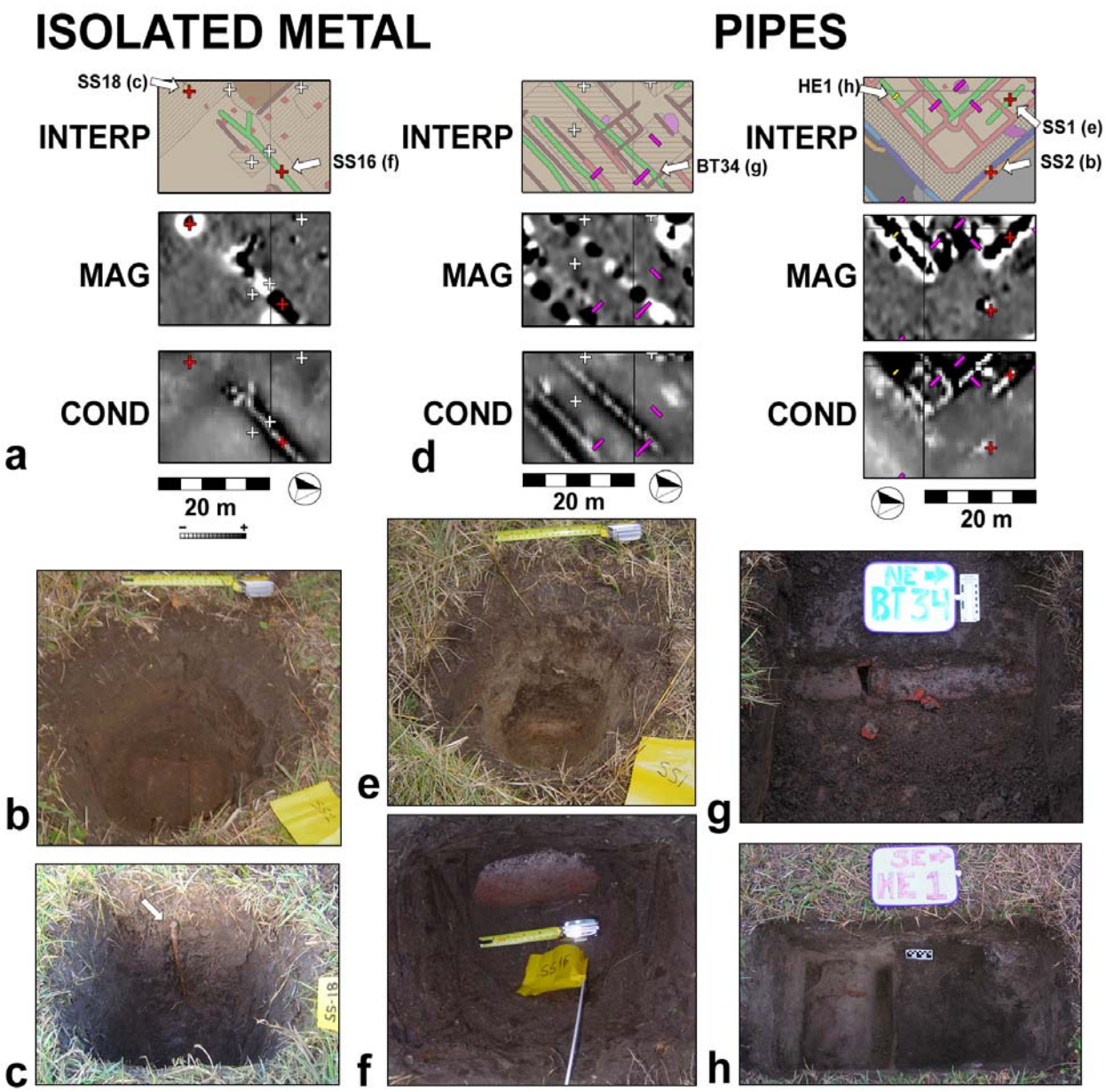


Figure 5.63. *Samples of archaeological sources for anomalies classified as "isolated metal" and "pipes." a) Vector interpretations, magnetic, and conductivity imagery showing loci of anomalies and excavations. b) STP SS2 showing an antique Kansas license plate. c) STP SS18 exposing a lightening rod. d) Vector interpretations, magnetic, and conductivity imagery showing locus of anomalies and excavations. e) SS1 revealing an iron pipe segment. f) SS16 showing a ceramic pipe segment. g) BT34 and a ceramic pipe segment. h) HE1 and a large steel sheet-metal pipe. KEY: INTERP=interpretations; MAG=magnetic gradiometry; COND=conductivity; BT=backhoe trench (purple); SS=single STP (red); HE=hand excavation (yellow).*

Pipes also represent a robust artifact type easily detected by several geophysical techniques. Unfortunately, in preparing names for the anomaly classes when they were first defined (Table 5.16), "pipe" anomalies were labeled as "metal pipes." This occurred

because large dipoles were characteristic in the magnetic gradiometry data as well as large "negatives" in the EM conductivity, both suggestive of iron (Figure 5.63d). This label also arose through general ignorance of materials employed in early 20th century sites. In the rush to classify the data for fieldwork it was assumed that pipes must be of iron (owing largely to the dipole signature), but this inference was also heavily influenced by a cast iron pipe excavated at the site by Larson and Penny (1997:58). Introspection and greater knowledge might also have suggested that fired ceramic pipes might also be present, as subsequent excavations revealed. Ironically, the archaeological evaluation team chose to zero in on the term "metal" in "metal pipe," whereas the SERDP evaluation team focused on the term "pipe," causing one source of difference in the accuracy of the evaluations. The archaeological team lumped all isolated metal artifacts and pipes into a single class whereas they are treated as individual types here. Figure 5.63e and Figure 5.63g both illustrate iron pipes exposed by excavation, while Figure 5.63f and Figure 5.63g indicate ceramic pipes. In general, STP were employed to test these relatively easy-to-locate targets; several were also exposed by backhoe (Figure 5.63g) and a hand excavation was employed to thoroughly document a single occurrence (Figure 5.63h). These classes of anomalies represent "high confidence" classes generally proven to represent correct targets in excavation (see below).

What were presumed to represent walls—linear anomalies meeting at right angles—were most clearly mapped using resistivity and GPR data (Figure 5.64a,b), but examples occurred in all geophysical data sets (Section 3.3.4.1). In general, all concrete walls in Army City are well revealed by resistivity survey as high contrast anomalies, while other walls represented largely by soil stains are better indicated by GPR. Anomalies interpreted as walls were categorized into two types: those robustly indicated (generally very large positive measurements and wide) and those only subtly revealed (with less extreme measurements). The latter were sometimes difficult to clearly discern in the data, but visualization of them was enhanced through extreme contrast manipulation. "Walls" were the most numerous of anomaly classes (Table 5.16) and most tested by excavation. Backhoe trench BT2 conveniently exposed adjacent robust walls of concrete in the Hippodrome (Figure 5.64c) and trench BT5 shows a heavily burned wall remnant of the Orpheum Theater (Figure 5.64d). "Robust walls" represent a high confidence class generally proven to represent correct targets in excavation (see below).

Subtle walls, however, were more difficult to define and locate in excavation, owing to "low contrast" excavation evidence. Many could be discerned only as subtle soil color, texture, or grain size changes with indistinct boundaries. BT6 shows a clear linear soil stain that quite possibly represents a former wall course (Figure 5.64e). Subtler evidence is exposed in BT38 where a few concrete or masonry fragments are seen in linear arrangement (Figure 5.64f). By far, Figure 5.64e,f represent the nature of the majority of wall anomalies investigated at Army City. "Subtle walls" represent a class of only moderate confidence representing correct targets in only a fraction of the excavations (see below).

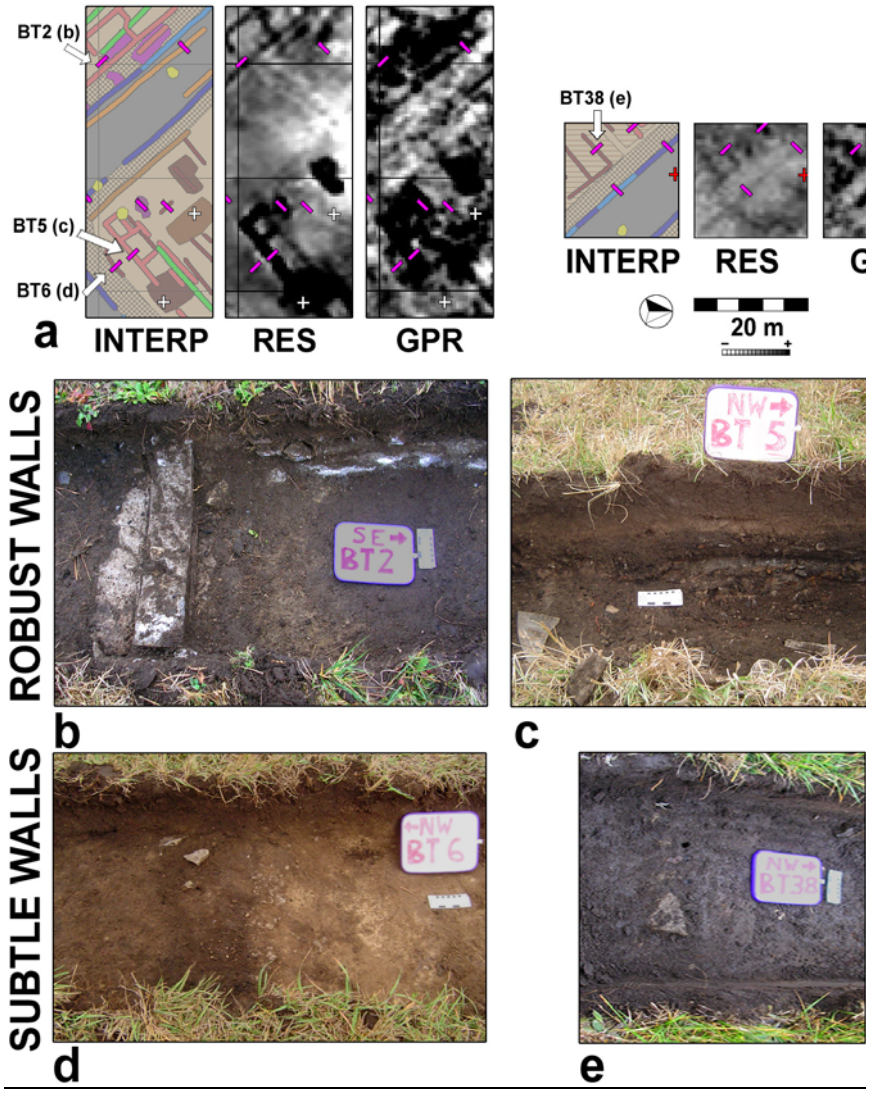


Figure 5.64. *Samples of archaeological sources for anomalies classified as "walls." a) Vector interpretations, resistivity, and GPR imagery showing loci of anomalies and excavations. b) BT2 showing concrete walls of the Hippodrome. c) BT5 showing a massively burned wall of the Orpheum Theater. d) Subtle wall in BT6 revealed as a soil stain. e) Subtle wall in BT38 defined by a line of concrete and masonry rubble. KEY: INTERP= interpretations; RES=resistivity; GPR=ground-penetrating radar; BT=backhoe trench (purple).*

Anomalies classified as "floors," in short-hand, were actually interpreted to represent anomalies of all types possessing area (as opposed to linear anomalies) that generally exhibited square to rectangular shapes. Most, indeed, likely represent the loci of former floors of the many structures within the town, but some might represent other types of features such as filled cellars, garden plots, rectangular spaces between buildings, and the like. These anomalies were also classified as robust versus subtle, with the former best defined by resistivity and GPR and the latter indicated by all methods. A sampling of such anomalies is illustrated in Figure 5.65a-c. For some robustly indicated anomalies sources were very easy to recognize in excavation. One exposed in the BT32 trench was composed of large pieces of concrete rubble (Figure 5.65a,d), while another in BT45 showed a layer of broken stone, concrete, and masonry (Figure 5.65e). Sources of even robustly classified floor anomalies were sometimes difficult to discern in excavation. Frequently, soil profiles had to be intensively examined to search for and define linear bands that might represent a former floor. This is indicated in (Figure

5.65f,g) where distinct 10 cm layers can be seen in profile (see arrows) that likely caused the area anomalies (Figure 5.65c).

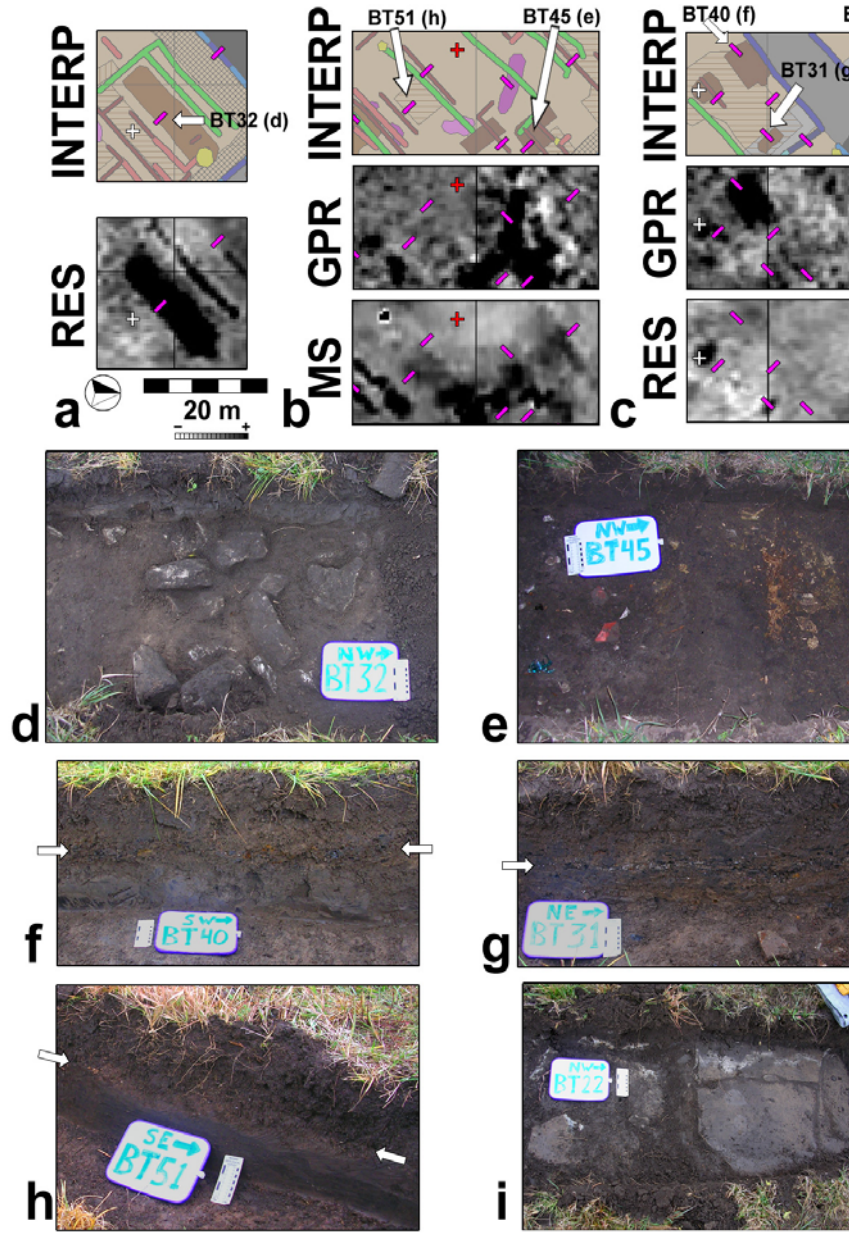


Figure 5.65. *Samples of archaeological sources for anomalies classified as "floors" and "debris piles." a-c) Vector interpretations, resistivity, GPR, and MS imagery showing loci of anomalies and excavations. d) BT32 showing concrete floor rubble. e) BT45 showing concrete and masonry rubble defining a floor. f) Robust wall anomaly seen in profile in BT40. g) Robust wall anomaly seen in profile in BT31. h) Subtle wall anomaly seen in profile in BT51. i) Concrete sidewalk slabs that caused a "debris pile" anomaly in BT22. KEY: INTERP= interpretations; RES=resistivity; GPR=ground-penetrating radar; MS=magnetic susceptibility; BT=backhoe trench (purple).*

Given the formation processes of Army City it is not difficult to understand why anomalies representing former floors may be difficult to validate by excavation. Shortly after the town was abandoned it was bulldozed, scraped and trucked away (Hargrave et al. 2002). Consequently, materials likely to generate high contrast anomalies, such as concrete, brick, and tile, were removed, leaving behind only a substrate of sparse remnant materials (e.g., basal sand or gravel) and perhaps compaction differences. It is therefore not surprising that, as a class, "floor or area/surface" anomalies were most difficult to validate (see below). It was particularly difficult to locate source evidence for floor

anomalies classified as "subtle." A case in point is a magnetic susceptibility anomaly that vaguely defines a rectangular space in Figure 5.65b, which was tested with trench BT51. It indicates a faint, but distinct, soil horizon of debris and gravel in profile (Figure 5.65h).

A related class of area anomaly was defined as the "debris pile" (Table 5.16). This anomaly type covers areas of irregular shape. They are generally very robust in appearance and tend to be revealed by resistivity, GPR, and magnetic susceptibility. They are inferred to represent concentrations of debris resulting from the dismantling of the town. One such occurrence is illustrated in BT22 that apparently represents an irregular polygon of remnant sidewalk slabs that somehow missed removal during the town's dismantling (Figure 5.65c,i). They generally represent a "high confidence" class pointing to correct targets in excavation (see below).

Street and alley gutters represent an unusual archaeological feature class, but a large fraction of the anomalies at Army City appear to correspond with the edges of the known street system (mapped in Figure 3.1). From photographs (Figure 3.2) it is obvious the streets were of bare earth and pronounced gutters appear on their edges, leading to the inference that the perimeter anomalies largely represent gutters and, in some case, remnant curbs. Plat maps indicate Army City had a series of streets and alleys, so a four-way division of anomalies was established based on robustness: robust and subtle street gutters and robust and subtle alley gutters. Many of the geophysical data sets give hints of these anomalies, but they are most consistent in magnetic susceptibility (probably because of deposition of magnetically enriched sediments) and GPR. Several are illustrated in Figure 5.66. Robust anomalies representing street gutters are illustrated by findings in trenches BT49 and BT26 (Figure 5.66c,d). They tend to be wide—more than a half-meter—and filled with robust debris, such as pieces of concrete, curb, roof shingles, and nails. Subtle anomalies pointing to street gutters are also highly visible archaeologically, but tend to be narrower in width. The one in BT39 is only 10 cm wide (Figure 5.66e), for example. Others may be defined by broken curb pieces (Figure 5.66f). It seems apparent from this evidence that the relative robustness of the corresponding geophysical anomalies arises principally from the width of the feature.

Anomalies associated with alley gutters are generally much subtler, as evidenced by their smaller and less contiguous expressions in Figure 5.66g. When exposed by excavation they generally are more difficult to see, but with a little effort they are made clear. The SERDP evaluation team soon learned that they could be better witnessed and mapped in profile, as illustrated in trench BT11 (Figure 5.66h). The ephemeral nature of a subtle alley gutter is indicated in BT10 where only a lens of pebbles filling the gutter shows its locus in cross-section (Figure 5.66c,i). All gutter anomaly types represent "high confidence" classes generally indicating correct targets in excavation (see below).

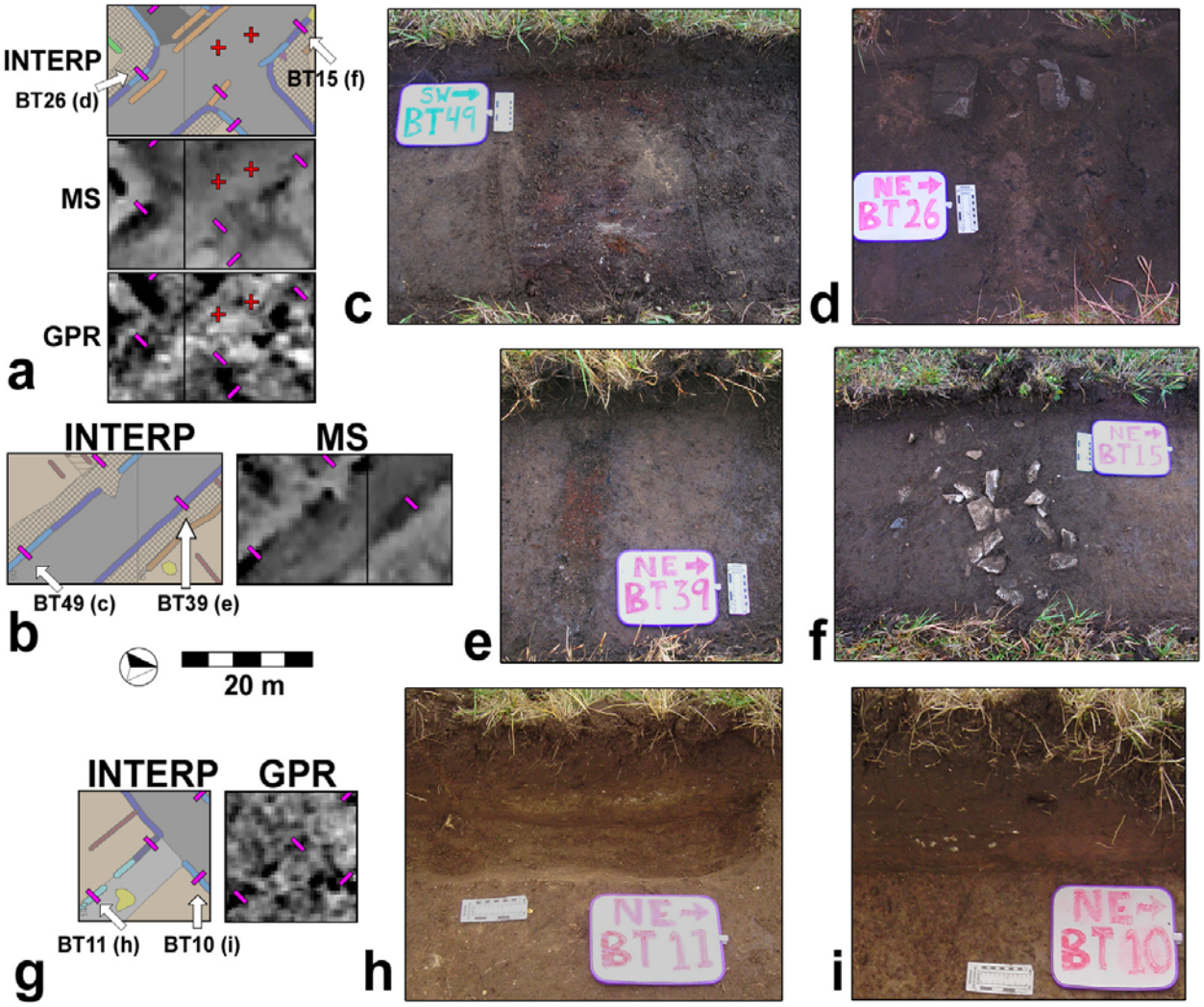


Figure 5.66. *Samples of archaeological sources for anomalies classified as street and alley "gutters." a,b) Vector interpretations, magnetic susceptibility, and GPR imagery showing loci of street anomalies and excavations. c) BT49 showing robust street gutter. d) BT49 showing robust street gutter with pieces of curb. e) Subtle street gutter in BT39. f) Subtle street gutter in BT15 with pieces of curb. g) Vector interpretation and GPR imagery showing loci of anomalies representing alley gutters and excavations. h) Robust alley gutter in BT11 profile. i) Subtle alley gutter shown by pebbles in BT10 profile. KEY: INTERP= interpretations; MS=magnetic susceptibility; GPR=ground-penetrating radar; BT=backhoe trench (purple).*

The final classes of culturally associated anomalies defined at Army City are also associated with its streets, and represent a collection of varied types. Very high resistivity in General Street near its intersection with Washington Avenue (Figure 5.67a) suggests some sort of covering, such as gravel or sand, forming a "street surface" class. Trench BT25 was placed in the center of this area and did, indeed, reveal a thick lens of sandy deposit 8-15 cm thick (Figure 5.67b,c). Several other trenches in west General Street showed similar deposits. Trench BT17 was placed to locate the source of an anomaly classified as a "street rut" (of which several can be seen in the historic photo of Figure 3.2). Anticipated to be a linear archaeological feature of different soil and fill, this

anomaly was deemed verified in trench BT17 (Figure 5.67d). The GPR data also reveal linear anomalies that parallel the street edges, but outside of the street (e.g., Figure 5.67a). They can only represent the inside edges of the sidewalks of which the town was so proud (seen in Figure 3.2). Speculation about sources for these anomalies suggested that drooping cement fragments might yet remain along the inside line of the walks. An alternate hypothesis is that walks might have been sanded in winter, and some of this sand may have accumulated along the inside walk edge. This later interpretation appears correct. In two backhoe trenches a small, linear lens of sand could be traced that paralleled the line of the former sidewalk, indicated by arrows in trench BT27 (Figure 5.67e).

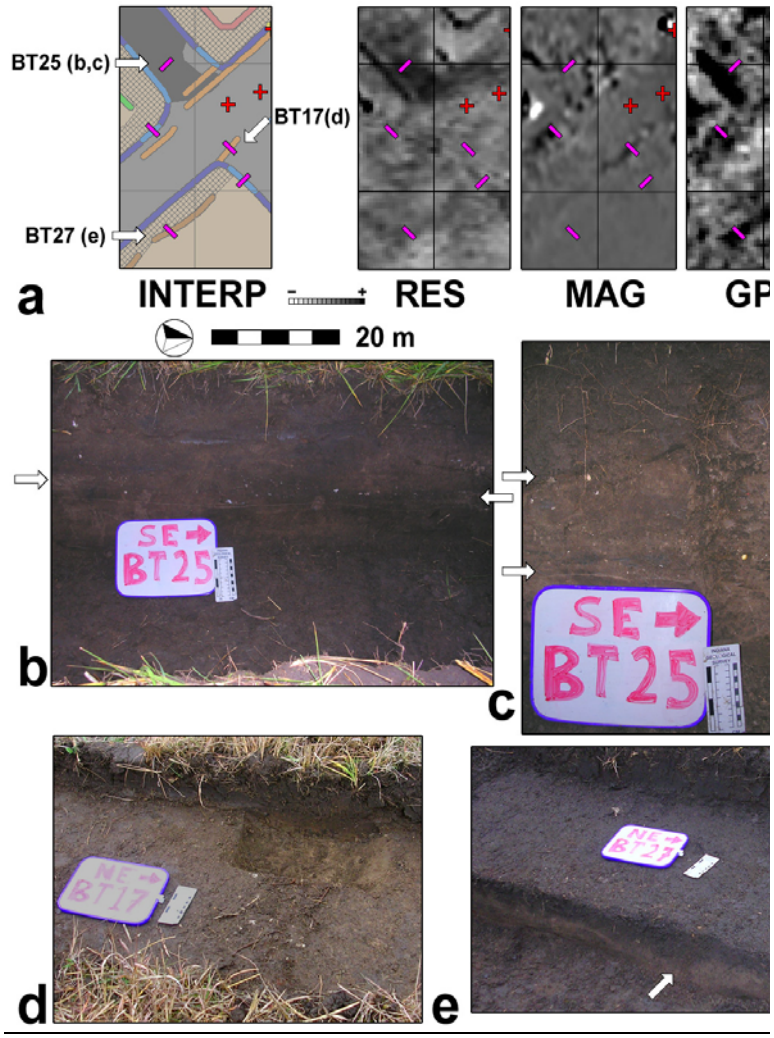


Figure 5.67. *Samples of archaeological sources for anomalies classified as "street surface," "street rut," and "sidewalk edge." a) Vector interpretations, resistivity, magnetic gradiometry, and GPR imagery showing loci of anomalies and excavations. b,c) BT25 showing sand lens (arrows) that may have formed a street surface. d) BT17 showing a linear street rut. e) Lens of sand in BT39 that may define the course of the former sidewalk edge. KEY: INTERP= interpretations; RES=resistivity; MS=magnetic susceptibility; GPR=ground-penetrating radar; BT=backhoe trench (purple).*

The ultimate class defined for archaeological testing represents the "normal" background and is formed by a collection of excavation samples (STP) without indications of cultural anomalies.

Accuracy assessment. The foregoing suggests good correspondence between archaeological feature types predicted by the combined remote sensing and actual

circumstances in the ground. The quantitative data support this view. In Table 5.28 the third column indicates the number of excavation units in which a source that explained the anomaly was identified. Discounting the "background" samples (which were sited over "normal" locations lacking anomalies), 78 (84%) of the 93 excavation units placed over likely anthropogenic anomalies contained in-ground elements that explained the anomalies. The forth column of Table 5.28 shows those excavations where archaeological features matched the predicted anomaly type, and the next three columns reports the "confidence" of the assessment. Seventy-one (76%) of the excavation units revealed targets of the correct predicted type.

Table 5.28. *SERDP team evaluation statistics for Army City's anomaly types. Empty cells represent zero counts.*

	Anomaly Class	Anomaly Source Identified?	Anomaly Type Correct?	*Confidence*			Target Distinct?	Iron present?	Number & Type of Excavations
				Low	Medium	High			
1.	Chimney base	1 (100%)	1 (100%)			1	1		1: 1 HE
2.	Footings	7 (100%)	6 (86%)			7	7		7: 6BT + 1HE
3.	Isolated metal	3 (100%)	3 (100%)			3	3	3	3: 3SS
4.	Pipes	5 (100%)	5 (100%)			5	5	5	5: 4SS + 1HE
5.	Debris pile	5 (100%)	5 (100%)		1	4	1	4	5: 5BT
6.	R wall	10 (100%)	9 (90%)			10	10	5	10: 10BT
7.	S wall	10 (71%)	8 (57%)	5	3	6	4	3	14: 14BT
8.	R floor	8 (67%)	7 (58%)	3	6	3	4	6	12:7BT + 5DS
9.	S floor	4 (40%)	2 (20%)	6	3	1		3	10:5BT + 5DS
10.	Street surfaced	1 (50%)	1 (50%)	1		1	1		2: 1BT + 1HE
11.	R street gutter	7 (100%)	7 (100%)			7	5	3	7: 7BT
12.	S street gutter	7 (100%)	7 (100%)	1	2	4	4		7: 7BT
13.	Street rut	2 (100%)	2 (100%)		1	1	1	1	2: 2BT
14.	R alley gutter	3 (100%)	3 (100%)			3	3	2	3: 3BT
15.	S alley gutter	3 (100%)	3 (100%)		1	2	1	1	3: 3BT
16.	Sidewalk edge	2 (100%)	2 (100%)		1	1	2		2: 2BT
17.	Background	11* (100%)	11* (100%)			11			11: 11SS
Totals (with #17):		**89 (86%)**	**82 (79%)**	**17 (16%)**	**18 (17%)**	**69 (66%)**	**52**	**36**	**104****
Totals (without #17):		**78 (84%)**	**71 (76%)**	**17 (18%)**	**18 (19%)**	**58 (62%)**	**52 (56%)**	**36 (39%)**	**93**

*This class represents the lack of anomalous indications and associated targets. **This number is less than the 106 of the planned sampling design because 2 HE units were paired with others for larger excavation exposures. KEY: R=robust; S=subtle; BT=backhoe trench; DS=double STP; HE=hand excavation; SS=single STP

Although sample sizes are far too small for meaningful statistical treatment of individual anomaly classes, the previous totals are associated with large sample numbers garnered from a random sampling program. Consequently, 95% confidence limits may be developed that indicate a physical source for an indicated anomaly is archaeologically locatable between 75-90% of the time. Additionally, between 66-84% of the indicated *types* of anomalies may be correct. Inspection of Table 5.28 indicates that two classes of anomalies were particularly problematic: subtle walls and floors. Removing them from consideration leaves 69 excavation units, with 93% showing an anomaly source and 95% confidence limits ranging between 84-97%. Similarly, 88% of the anomaly types are now

correct with 95% confidence limits from 79-94%. It is emphasized that these statistics may only bear on Army City's unique combination of archaeological features, soils, and environmental conditions.

Various patterns in the data are of interest (Table 5.28). Most of the anomaly classes showed high accuracy in the evaluation. Robust walls were often of concrete, the chimney base was made of stone and concrete, and footings were of concrete and brick, all targets that yield high contrast anomalies and near-perfect sensing accuracy. The same is also true of metal artifacts and pipes (Table 5.28) that give unambiguous geophysical signatures. Surprisingly, given their frequently small size and ephemeral character as simple soil stains, all predicted gutters (Figure 5.66), whether in streets or alleys, were validated by excavation. Subtle walls, floors of all types, and the "street surface" classes suffered loss of accuracy, however. This is not surprising for these area anomalies. As mentioned earlier, few exhibited hard floor surfaces and most could be witnessed only by subtle stains and through detailed examination of profiles. The project may have been overly ambitious in defining linear anomalies as walls, as this was another source of error (Table 5.28).

The nature of the excavation unit employed has a bearing on accuracy. A hand excavation (HE) that tried to define a street surface opened too small an area for effective evaluation. STP were also too small to confidently discern floors in limited exposures. It is difficult to compare different modes of excavation quantitatively. For most feature types different excavation strategies were employed. Although hand excavations were applied to investigate several archaeological feature types in detail, their numbers are far to low for quantitative treatment.

Table 5.29. *Comparison of accuracy rates between the limited exposures of the STP (DS) and the backhoe trench (BT).*

Subtle Floors			
Type of Excavation	**Anomaly Source Identified?**	**Assigned Anomaly Type Correct?**	**Total Samples**
DS	1	0	5
BT	3	2	5
Robust Floors			
Type of Excavation	**Anomaly Source Identified?**	**Assigned Anomaly Type Correct?**	**Total Samples**
DS	3	2	5
BT	5	5	7

KEY: BT=backhoe trench; DS=double STP

Two types of excavation were applied to investigate anomalies classified as floors, however: the backhoe trench and the double STP. The latter was utilized under the mistaken premise that intact and "solid" floors might yet remain in much of the site where an STP would "hit" a distinct surface (no intact floors were encountered in any unit). In fact, much of the error in Table 5.28 derives from use of the STP in this context. This is illustrated in Table 5.29 where a comparison of accuracy rates is made for floor anomalies between the limited exposures of the STP and the wide expanse of the backhoe trench. Clearly, the wider exposure allowed inspection of soil changes and the general

visibility of more evidence allowing substantially improved accuracy of the assessments. Backhoe trenches outperformed STP for locating anomaly sources and in identifying whether the anomaly was interpreted as the correct type.

Fundamentally, whether or not an archaeological source for an anomaly is recognized depends in large part on its distinctiveness. Vague stains with indistinct boundaries are less likely to be recognized as archaeological features than ones with sharp, well-defined boundaries (e.g., some of the gutters in Figure 5.66). This characteristic of field testing is illustrated in Table 5.30, which charts the "confidence" placed in the assessment of archaeological feature types exposed in the excavations against their distinctiveness. It is apparent that distinct archaeological features, with sharp edges and well-defined boundaries, are more confidently assessed.

Table 5.30. *Distinctly defined archaeological features with sharp boundaries are more likely to be confidently assessed.*

Target Distinct?	**Confidence of Assessment**			**Totals**
	Low	***Medium***	***High***	
No	16	16	9	41
Yes	1	2	49	52
Totals:	17	18	58	93

Effectiveness of individual geophysical methods. Sorting out which specific geophysical sensors best defined anomalies that could be reliably interpreted is particularly difficult at Army City because so many of the interpretations were based on integrated data and so many sensors were redundant for many kinds of features. For example, pipes were strongly indicated by EM conductivity and magnetic gradiometry; many walls were simultaneously revealed by resistivity and GPR; street gutters were generally defined in parallel by magnetic susceptibility and GPR. For each of the excavation units of Table 5.28 the geophysical data set that gave the *most* robust indication of the associated anomaly was assigned to that unit. As noted, this made for a difficult assessment because many archaeological features were simultaneously and strongly revealed by several sensors.

The data (Table 5.31) indicate that of all the techniques, magnetic gradiometry is most accurate with 100% of the indicated anomalies associated with a source and 100% classified to the correct type. It is emphasized that nearly all of the magnetic gradiometry anomaly types point to metallic artifacts or fired ceramic pipes associated with distinctive anomalies and easy to locate in the ground. Most of the anomalies defined at Army City were most visible in the GPR data. It is the next most reliable, with 97% of the excavations revealing an anomaly source and 90% accurately classified to type. Magnetic susceptibility and resistivity both defined an equal number of anomalies (Table 5.31) and both show moderate error, with resistivity performing better. Most of the error is associated with the ephemeral floor problem, discussed earlier. EM conductivity and thermal infrared, although indicating many strong anomalies (Section 4.6.1), where generally outperformed by other sensors that often revealed the same anomalies more strongly. These data types show little impact in Table 5.31.

5.16.2. Field Investigations at Pueblo Escondido

Michael L. Hargrave, Construction Engineers Research Laboratories (CERL), and Eileen G. Ernenwein, University of Arkansas

Excavations at Pueblo Escondido were conducted by TRC Environmental Corporation's El Paso Texas office (hereafter referred to as "TRC"). Co-field directors for TRC were Mr. Paul Lukowski and Ms. Elia Perez, assisted by a crew of five. TRC was selected to conduct this work because of their extensive experience with the local archaeology and material culture, and their outstanding history of work at Fort Bliss.

Table 5.31. *Anomaly types by the principal geophysical device that defined them.*

Anomaly Type		COND			GPR			MAG			MS			RES			THERM		
		Tot	Src	Type	Tot	Src	Type	Tot	Src	Type	Tot	Src	Type	Tot	Src	Type	Tot	Src	Type
1.	Chimney base													1	1	1			
2.	Footings													7	7	6			
3.	Isolated metal							3	3	3									
4.	Pipes							5	5	5									
5.	Debris pile				4	4	4							1	1	1			
6.	R wall				4	4	3							7	6	6			
7.	S wall				8	7	6				3	1	0	3	2	1			
8.	R floor				3	3	3				2	2	2	5	3	2	1	0	0
9.	S floor	1	0	0							9	4	2						
10.	Street surfaced													2	1	1			
11.	R street gutter				4	4	4				3	3	3						
12.	S street gutter				1	1	1				6	6	6						
13.	Street rut				1	1	1	1	1	1									
14.	R alley gutter				2	2	2				1	1	1						
15.	S alley gutter				1	1	1				2	2	2						
16.	Sidewalk edge				2	2	2												
	Totals:	**1**	**0**	**0**	**30**	**29**	**27**	**9**	**9**	**9**	**26**	**19**	**16**	**26**	**21**	**18**	**1**	**0**	**0**
	(%)		0%	0%		97%	90%		100%	100%		73%	62%		81%	69%		0%	0%

KEY: COND=conductivity; GPR=ground-penetrating radar; MAG=magnetic gradiometry; MS=magnetic susceptibility; RES=resistivity; THERM=thermal infrared; R=robust; S=subtle; Tot=total excavation units; Src=anomaly source located; Type=anomaly classified to correct archaeological type.

The SERDP team was represented by Kvamme, Ernenwein, and Hargrave, with assistance from two University of Arkansas graduate students (Goodmaster and Markussen). TRC's investigations occurred from 25-29 January 2005. The SERDP team

arrived at the site several days before TRC in order to place flags that marked the locations for excavation units (Section 5.15.3). Field conditions ranged from good to difficult, with rain and a brief hailstorm on 27 and 28 January.

The geophysical survey area at Pueblo Escondido consisted of twenty-five 20 by 20 m blocks that covered a total area of 10,000 m^2. Geophysical anomalies were assigned to both archaeological and descriptive categories (Section 5.12.5). The anomaly types included the following: room floor, magnetic room floor, inferred room floor, outlying floor, burned room wall, robust room wall, subtle room wall, interior room feature, magnetic feature in room, robust lineation, subtle lineation, small systematic feature, small random feature, and small magnetic feature. Adjectives such as subtle and robust were used to describe the level of contrast displayed by the anomalies, which was generally related to feature characteristics (size, density, 'massiveness') and depth. Random features were spatially isolated whereas systematic features appeared to be part of a distributional pattern of similar anomalies.

Although the original sampling design stipulated 123 excavation units (including mechanized trenches, shovel scrapes, and small hand-dug units) and 80 cores (Table 5.21), it was modified during fieldwork. Initial plans had called for the excavation of 26 small (.5 by .5 m) hand-excavation units and 60 shovel scrapes, each measuring 2 x .25 m. Both of these excavation categories had to be abandoned in response to site conditions. Cultural deposits proved to extend much deeper (up to 60 cm below present ground surface) than had been predicted based on reports of previous excavations (Beckes 1975; Hedrick 1967), particularly in the northwestern portion of the study area. Shovel scrapes and hand excavations proved too time-consuming, so the effort was limited to mechanized trenches and cores. Several of the shovel scrapes and hand excavations were replaced with mechanized trenches to keep testing of each stratum proportional to the original stratified random sample. This led to some minor problems when testing small anomalies with mechanized trenches, as some of them may have been removed with the backhoe before they were noticed, with little or no trace left behind in the trench walls. Poor weather also diminished available field time. The excavations ultimately included 1 hand excavation unit, 3 shovel scrapes, 28 auger tests, and 43 backhoe trenches for a total of 74 tests of anomalies (Figure 5.68). Three-inch diameter auger tests were excavated in 10-15 cm levels and all soil was screened through 1/8-inch mesh. Ten pairs of auger tests (that is, a total of 20 holes) were excavated. In each pair, one hole was located inside the anomaly and the other was placed outside in order to provide a basis for comparison. The remaining 8 auger tests were individually excavated and targeted on anomalies. Three other single auger tests sampled "background" areas where no anomalies were detected. In addition to the 3-inch auger, a 1-inch Oakfield soil probe was also used to document the depth and fill characteristics of various deposits (TRC Env. 2005:31). One mechanized trench and three cores (1.5 pairs) were excluded from the analysis because not enough information was uncovered to verify the presence or absence of the anomaly's source. Table 5.32 summarizes the final distribution of test units according to anomaly types.

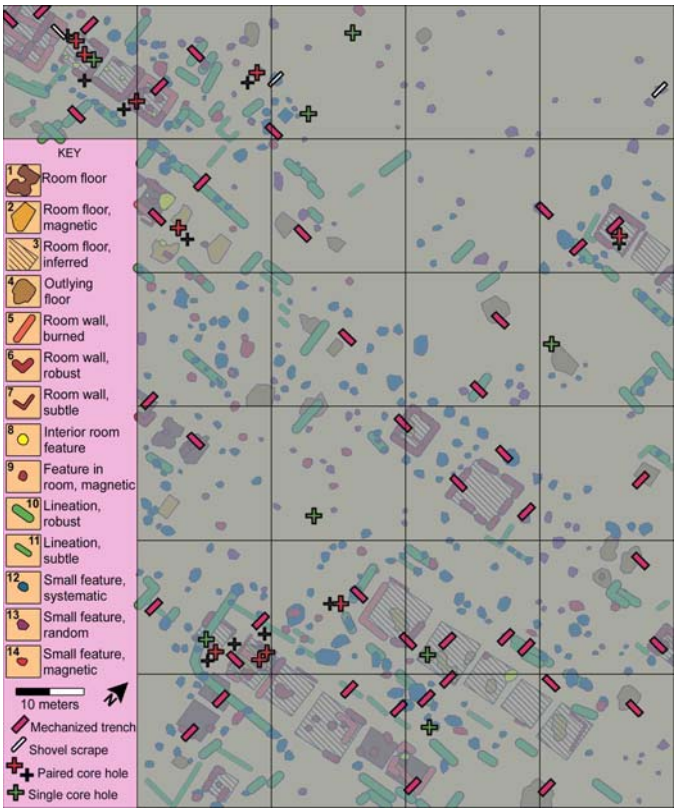


Figure 5.68. *Final distribution of units for Pueblo Escondido test excavations overlaid on the vector interpretation. Linear anomalies and excavation unit symbols are exaggerated.*

Trenches were excavated using a backhoe equipped with a 75 cm wide toothless bucket. Particular care was taken to prevent the backhoe from running across suspected architectural features (TRC Env. 2005:31). Soil excavated from trenches was placed in discrete piles to minimize impacts to surface artifact distributions. The backhoe excavated in a series of very shallow (5-10 cm) cuts, and excavation ceased when it was clear that a feature was present. All excavations were documented by scaled maps, digital photographs, detailed notes, and standardized forms.

Table 5.32. *Actual number and distribution of units excavated at Pueblo Escondido.*

Code	Anomaly type	Stratum	Population size (# of anomalies , total area* in m^2)	Test unit types				Total samples
				M	SS	P-C	S-C	
1	*Room block floor*	R	12, 150.6	3		2		**5**
2	*Magnetic room block floor*	R	14, 90.9	3			1	**4**
3	*Inferred room block floor*	R	20, 637.7	1		1*		**2**
4	*Outlying floor*	R	21, 147.3	5				**5**
5	*Magnetic room block wall*	R	21, 86.3	2	1	1		**4**
6	*Robust room block wall*	R	34, 122.0	4		1		**5**
7	*Subtle room block wall*	R	17, 51.2	2				**2**
8	*Interior room block feature*	R	16, 16.2	1		1		**2**
9	*Circular magnetic interior room block feature*	R	27,32.6	1		1	2	**4**
			Totals: 182, 1334.8	**22-1x2m**	**1**	**6**	**3**	**33 (46%)**
10	*Robust lineation*	L	72, 255.8	6				**6**
11	*Subtle lineation*	L	31, 103.2	3				**3**
			Totals: 103, 359.0	**9-1x3m**				**9 (13%)**
12	*Small polygon systematically distributed*	P	180, 336.5	7	1			**8**
13	*Small polygon randomly distributed*	P	93, 83.8	4	1			**5**
14	*Circular magnetic exterior anomaly*	P	46, 65.4		0	2	2	**4**
			Totals: 319, 485.7	**11-1x2m**	**2**	**2**	**2**	**17 (24%)**
15	*Background*	B	∞, 7820.5			8	3	**11 (17%)**
			Totals: 604, 2179.5	**42 (71m)**	**3**	**17**	**8**	**70**

Abbreviations: R=Room block; L=Major lineation; P=Point/polygon anomalies; B=background; M=mechanized 1x2m or 1x3m trench; S-S =shovel scrape; P-C=paired core; S-C=single core.
*Area for linear classes calculated as 1 linear meter = 1 square meter.

As was the practice at each site, the SERDP team assisted in monitoring the backhoe excavation but focused primarily on an independent documentation of each

excavation unit (Figure 5.69). Here the focus was on identifying subtle soil characteristics that might explain the existence of anomalies, particularly in situations where no obvious archaeological cause was present.

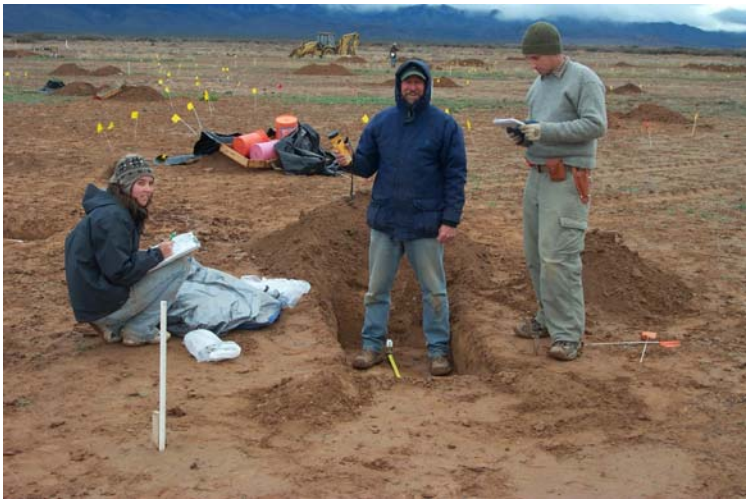


Figure 5.69. *SERDP team members using a KT-9 Kappameter hand-held magnetic susceptibility meter while documenting a backhoe trench at Escondida Pueblo, January 2005.*

Although artifacts were very abundant on the surface (probably as a result of wind erosion), they were very sparse in nearly all subsurface contexts. Soil strata were very difficult to discern. The TRC report observes that "The deposits often contained scattered charcoal and intermixed sediment fill of varying consistency (e.g., adjacent patches of silty sands with differing clay content that could be interpreted as mixed areas of erosional in-filling and wall/room fall) that required time consuming fieldwork in order to make positive identification of the expected anomalies" (TRC Env. 2005:31).

5.16.2.1. Evaluation of predictions by archaeological team
Michael L. Hargrave, CERL

Tables 5.33 and 5.34 show the results of ground truthing at Pueblo Escondido for the backhoe trenches and auger tests (respectively). If results of the two excavation techniques are combined, one sees that 30 of the 61 excavation units (49.2%) verified the presence of the type of feature that was predicted based on the geophysical data. In only one case was some other type of feature present. No feature was visible in 37.7% of the cases (n=15), and the remaining 11.5% were indeterminate (n=7). Table 5.35 identifies the type of deposit documented for the various anomaly types.

Table 5.33. *Results of ground truthing for backhoe trenches at Pueblo Escondido (adapted from TRC Env. 2005:Table 7.1).*

Feature Type	Predicted Feature Type Present		Other Feature Type Present		No Feature Present		Indeterminate		Total
	N	%	N	%	N	%	N	%	N
Room floor	3	75.0%	0	0.0%	0	0.0%	1	25.0%	4
Room floor, magnetic	3	100.0%	0	0.0%	0	0.0%	0	0.0%	3
Room floor, inferred	1	100.0%	0	0.0%	0	0.0%	0	0.0%	1
Room floor, outlying	3	60.0%	0	0.0%	0	0.0%	2	40.0%	5
Room wall, burned	2	100.0%	0	0.0%	0	0.0%	0	0.0%	2
Room wall, robust	2	50.0%	0	0.0%	2	50.0%	0	0.0%	4
Room wall, subtle	2	100.0%	0	0.0%	0	0.0%	0	0.0%	2
Room interior feature	0	0.0%	1	100.0%	0	0.0%	0	0.0%	1
Room feature, magnetic	1	100.0%	0	0.0%	0	0.0%	0	0.0%	1
Lineation, robust	0	0.0%	0	0.0%	5	83.3%	1	16.7%	6
Lineation, subtle	0	0.0%	0	0.0%	2	66.7%	1	33.3%	3
Small feature, systematic	3	37.5%	0	0.0%	3	37.5%	2	25.0%	8
Small feature, random	0	0.0%	0	0.0%	3	100.0%	0	0.0%	3
Small feature, magnetic	0	0.0%	0	0.0%	0	0.0%	0	0.0%	0
Total	20	46.5%	1	100.0%	15	34.9%	7	16.3%	43

Table 5.34. *Results of ground truthing for auger tests at Pueblo Escondido (adapted from TRC Env. 2005:Table 7.2). Three other auger tests were targeted on background locations and no features were identified.*

Feature Type	Predicted Feature Type Present		Other Feature Type Present		No Feature Present		Indeterminate		Total
	N	%	N	%	N	%	N	%	N
Room floor	1	50.0%	0	0.0%	1	0.0%	0	0.0%	2
Room floor, magnetic	1	100.0%	0	0.0%	0	0.0%	0	0.0%	1
Room floor, inferred	0	0.0%	0	0.0%	0	0.0%	0	0.0%	0
Room floor, outlying	0	0.0%	0	0.0%	0	0.0%	0	0.0%	0
Room wall, burned	1	50.0%	0	0.0%	1	0.0%	0	0.0%	2
Room wall, robust	1	100.0%	0	0.0%	0	0.0%	0	0.0%	1
Room wall, subtle	0	0.0%	0	0.0%	0	0.0%	0	0.0%	0
Room interior feature	2	50.0%	0	100.0%	2	0.0%	0	0.0%	4
Room feature, magnetic	0	0.0%	0	0.0%	2	0.0%	0	0.0%	2
Lineation, robust	0	0.0%	0	0.0%	0	0.0%	0	0.0%	0
Lineation, subtle	0	0.0%	0	0.0%	0	0.0%	0	0.0%	0
Small feature, systematic	1	100.0%	0	0.0%	0	0.0%	0	0.0%	1
Small feature, random	0	0.0%	0	0.0%	1	100.0%	0	0.0%	1
Small feature, magnetic	3	0.0%	0	0.0%	1	0.0%	0	0.0%	4
Total	10	55.6%	0	100.0%	8	44.4%	0	0.0%	18

Table 5.35. *Results of ground truthing for backhoe trenches at Pueblo Escondido (adapted from TRC Env. 2005:Table 7.1).*

Feature Type Present:		Fill, floor	Fill, floor, wall	Feature	Nothing	Ambiguous	Total
Anomaly Category:							
Room floor	n	4			1		5
	%	80.0%			20.0%		100.0%
Room floor, magnetic	n	3				1	4
	%	75.0%				25.0%	100.0%
Room floor, inferred	n	1					1
	%	100.0%					100.0%
Room floor, outlying	n	2				2	4
	%	50.0%				50.0%	100.0%
Room wall, burned	n		2				2
	%		100.0%				100.0%
Room wall, robust	n	2			1		3
	%	66.7%			33.3%		100.0%
Room wall, subtle	n	1	1				2
	%	50.0%	50.0%				100.0%
Room interior feature	n	1					1
	%	100.0%					100.0%
Room feature, magnetic	n			1			1
	%			100.0%			100.0%
Lineation, robust	n				5	1	6
	%				83.3%	16.7%	100.0%
Lineation, subtle	n				2	1	3
	%				66.7%	33.3%	100.0%
Small feature, systematic	n			3	3	2	8
	%			37.5%	37.5%	25.0%	100.0%
Small feature, random	n				3		3
	%				100.0%		100.0%
Small feature, magnetic	n						
	%						
TOTAL	n	14	3	4	15	7	43
	%	32.6%	6.9%	9.3%	34.9%	16.3%	100.0%

TRC Env. (2005:71) reported that it was more difficult to make decisions about the presence or absence and nature of features in the auger tests than in the backhoe trenches. The trenches exposed much larger areas, providing far better opportunities to detect the subtle visual and texture contrasts that indicated the presence of some of the features. The excavators noted that differences in soil texture and compaction were better indications of features than were differences in color. Although carbon was relatively abundant at the site, it apparently had not broken down sufficiently to impart a generalized stain to many of the cultural deposits (TRC Env. 2005:71). The 3-inch auger churned the soil recovered from each level, making it almost impossible to detect subtle changes in texture and compaction. The auger holes were, of course, too small to permit any visual examination of their profiles. The most useful information from the augers was the recovery of carbon, artifacts, burned earth, or adobe, all of which suggested the possible presence of a feature. TRC did find that the use of paired augers, one excavated inside the anomaly and the other located outside, was a useful approach (Figure 5.70).

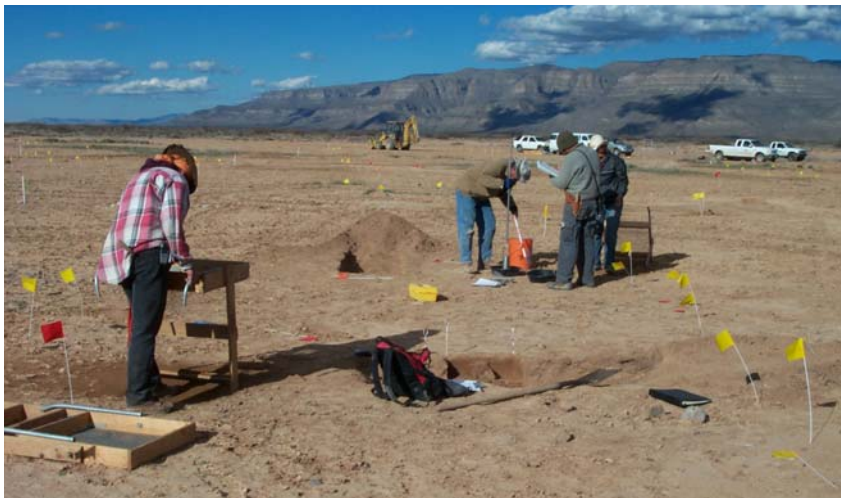


Figure 5.70. *TRC crew uses a 3-inch auger to investigate anomalies at Pueblo Escondido, January 2005.*

Contrary to these impressions about the usefulness of the two excavation techniques, the presence of the predicted type of feature was actually verified by auger tests a little more frequently (55.6%) than by backhoe trenches (46.5%) (Tables 5.33 and 5.34). Furthermore, 16.3% of the backhoe trenches resulted in indeterminate results concerning the presence of the predicted type of feature, whereas none of the auger tests were recorded as indeterminate. These results probably reflect a difference in the kind of data needed to make a decision about the presence of a feature. Because the trenches provided a far better opportunity to determine if a feature was present, the excavators probably demanded more convincing evidence. If such information was not available, they recorded the result as indeterminate. In contrast, the auger tests did not provide many options for making a decision. The presence of artifacts, carbon, burned soil, ash, adobe, or other discernable feature fill in an auger strongly suggested that a feature was present. The absence of such evidence required the excavators to record negative results. The use of this presence/absence criterion explains the lack of indeterminate results in the auger testing.

TRC found clear differences in the rates of verification of generalized anomaly categories. Features associated with structures were far more reliably predicted than were two categories of extramural (located outside of structures) features (Figure 5.71). In the backhoe trenches, 76.2% (16 of 21) of the predicted structural features (room floors, magnetic room floors, inferred room floors, outlying room floors, burned room walls, robust room walls, and subtle room walls) were verified. Room interior and magnetic interior room features are omitted here because they do not represent relatively massive structural components like floors and walls. In contrast, only 15% (3 of 20) of the non-structure features (robust lineations, subtle lineations, systematic small features, and random small features) were verified in trenches. It is perhaps not too surprising that relatively massive features like floors and walls are so reliably discernable in the geophysical data (although the value of this for archaeologists working in the region should not be under appreciated). It was perplexing, however, to see such a low rate of verification for the extramural anomalies.

Figure 5.71. *The sharp edge of a pit house with relatively rich fill in Trench 2 at Pueblo Escondido, January 2005.*

The anomaly categories investigated in the field plotted systematic small features (labeled as 12 in the key) in dark blue, and lineations (categories 10-11) in green (Figure 5.45). The lineations and the aligned systematic small features are oriented either parallel or perpendicular to many obvious structures. The systematic small features tend to be located close to the structures, and in some cases appear to completely surround individual rooms. These extra-mural anomalies are clearly not randomly distributed clutter. Yet very few of the lineation and small feature anomalies that were investigated were found to be associated with a feature. Three of the systematic small features that were verified were difficult to interpret in functional terms, although all three provided evidence of burning. Two of the features were moderate-sized (70 and 85 cm diameter) areas of scattered charcoal, and one was a small (15 cm diameter) concentration of ash (TRC Env. 2005:59-62). One possibility is that the extra-mural anomalies are, as a group, associated with highly ephemeral features. It may be that the three that were verified were simply relatively substantial or well-preserved examples. It would be logical to assume that many activities would occur outdoors and might well be located near the houses. Like the ruts in Army City's dirt streets, certain categories of ephemeral features at Pueblo Escondido might simply be discernable in the geophysical data yet exceedingly difficult to identify in a field situation, particularly at a site characterized by little staining and highly weathered deposits.

A second interpretation is that the extra-mural anomalies are associated with plant roots. It is possible that the subsurface structural remains retain a disproportionate amount of moisture and that this causes vegetation to grow more densely around structure perimeters. No evidence for this patterning was noticed in the field. However, Figures 6.4 and 7.62 appear to show several alignments of dense vegetation that are oriented roughly north-south. The aligned rooms visible in the geophysical data are oriented east-west, but it is conceivable that the short room blocks occur in north-south bands.

Although most anomalies interpreted as structural features such as floors and walls were verified by excavation, several were not. Examples include robust room block wall anomalies investigated by Trenches 23 and 37, a room block floor anomaly investigated by Trench 28, and the outlying floor anomalies in Trenches 27 and 34. TRC noted that a slight misplacement of Trench 23 could explain the absence of a feature, but such an error is viewed here as unlikely (TRC Env. 2005:63). The SERDP team was very aware of the importance of locating excavation units where they would be most likely to intersect features suspected to be associated with anomalies, and considerable care was taken to avoid location errors in the field. Trench 37 was positioned to intersect the western wall and floor of an apparently poorly preserved room. No artifacts were encountered, and TRC reported no indications of a feature (although a thin lens of sandy silt loam is shown on the trench profile map [TRC Env. 2005:100, Figure 10]). It is unlikely that the anomaly investigated by Trench 37 was misinterpreted, since it was one of several aligned rooms.

Trench 28 encountered one sherd, one flake, and scattered charcoal in the unit floor and profile, but no evidence for the predicted floor. However, a nearby trench produced much better indications of an adjacent room in the same alignment. Verification of the adjacent room provides strong support that the anomaly tested by Trench 28 was also associated with a room (TRC Env. 2005:64).

Results from Trenches 27 and 34, both excavated to investigate anomalies thought to be associated with outlying floors, were also ambiguous (TRC Env. 2005:72, Table 7.1). Trench 27 was less thoroughly documented than most other trenches, but the occurrence of artifacts suggests that a structure was probably present (TRC Env. 2005:64). Trench 34 was fully documented and encountered relatively abundant artifacts and charcoal. TRC concluded that the anomaly was associated with a trash pit "or possibly pit room structure fill" (TRC Env. 2005:65).

TRC was requested to make objective assessments about the deposits encountered, and this review of their characterization of the results for Trenches 27, 28, and 34 as ambiguous is not intended as a criticism. It is worth noting that, if the results of these three units were tabulated as positive rather than ambiguous, the verification rate for anomalies related to structures would be even better (90.5%) than the current 76.2%.

Trenches 23 and 37 are particularly important because they appear to provide evidence for the existence of cultural deposits that can be detected by geophysics but which are essentially "invisible" to standard excavation techniques. This characteristic may be due to several factors, including the nature of feature construction, use, and abandonment, and possibly the effects of looting. This possibility may be a bit disturbing to some archaeologists, but it is entirely plausible at a site characterized by little organic staining and highly weathered deposits.

5.16.2.2. Evaluation of predictions by SERDP team
Eileen G. Ernenwein, University of Arkansas

These evaluations of each excavation unit were made when fieldwork was completed utilizing a preliminary report from TRC Env. (2005) and field notes taken by the SERDP group on standardized forms. Comments made by TRC provided detailed information about what was encountered during excavation, the types of artifacts that

were collected, and interpretation of the findings from a local archaeological perspective. Information recorded on SERDP forms provided interpretations of the fully excavated units from a geophysical perspective, including detailed notes on soil stratigraphy and other phenomena that were sometimes overlooked by the local archaeologists. Both sources of information were used to make a final assessment of the accuracy of the SERDP project predictions.

Even though the Pueblo Escondido geophysical data were exceptionally detailed, predictions were made very carefully with specific labeling of features avoided. Clearly rectangular and linear anomalies were (mostly) predicted to be structures and walls, while some linear anomalies were simply predicted as "linear features" because of lack of certainty of what they might represent (Section 5.12.5). In addition, the myriad small, nondescript anomalies were predicted only to be "small features," and were grouped into categories based on their location relative to the predicted structures (interior versus exterior) and their spatial pattern (distributed randomly or systematically). Excavations were anticipated to reveal more specific archaeological feature categories; e.g. small, random features would turn out to be hearths and storage pits. Yet test excavations were optimized for verification of feature presence, so feature identity was rarely determined. There were some cases where excavations revealed very specific features, such as adobe versus pithouse walls. Nonetheless, the final interpretation (Section 5.17) makes use of fewer feature classes than the original prediction.

The fourteen (predicted) anomaly classes (Table 5.32) can be grouped into five general categories: floors (classes 1-4), walls (5-7), interior features (8-9), lineations (10-11), and small features (12-14). Of the fourteen floor anomalies tested with mechanized trenches and cores, thirteen were deemed "accurate" by the SERDP crew, making this the most reliable type of anomaly. Figure 5.72 illustrates a case where a burned or magnetically enhanced floor was predicted and confirmed by excavation. The outline of a structure was clearly indicated in GPR slices (Figure 5.72a), part of the interior area was anomalous in magnetic susceptibility (Figure 5.72b), and a smaller portion had a high magnetic gradient (Figure 5.72c). The test excavation, unit M16, revealed a distinct, well intact five-centimeter-thick clay or adobe floor buried .15 m below the surface. On top and within the floor were scattered pieces of charcoal, fire-cracked rock, and burned pottery sherds, which most likely contributed to the higher magnetic susceptibility and magnetic gradient readings. The test excavation confirms the floor prediction and shows that it was made of clay or adobe. The majority of the other floor anomalies were also verified, but most were compacted sandy loam rather than clay.

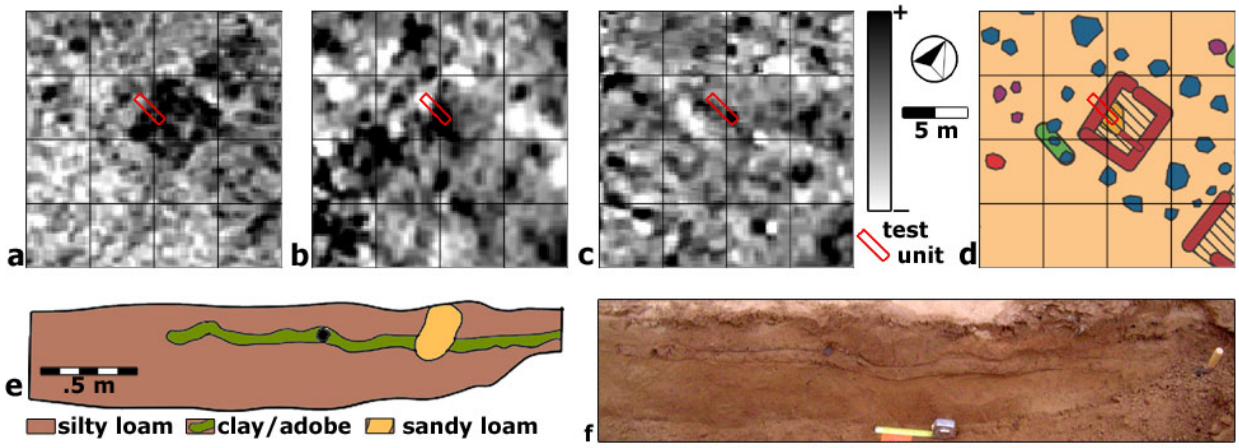


Figure 5.72. *Testing of a floor anomaly with mechanized trench M16. Small subareas surrounding the test unit are shown: a) GPR slice from .15-.34 m, b) magnetic susceptibility, c) magnetic gradiometry, and d) vector interpretation (for symbol key see Figure 5.68). e) Profile of the test unit's north wall (adapted from Lukowski & Perez 2005). f) Photograph of eastern portion of the test unit's north wall.*

Eleven wall anomalies were tested with one shovel scrape, two cores, and eight mechanized trenches. Only one trench showed no evidence, while another revealed a soil disturbance and a third turned out to be a burned wooden (probably roof) beam. The remaining eight were confirmed to be walls of one sort or another. Three of the predicted "walls" turned out to be the edge of pit-room basins, such as that shown in Figure 5.73, and were therefore deemed to be consistent with the predictions. Mechanized trench M2 was placed across the northern edge of the largest and most visible structure anomaly in the entire dataset (Figure 5.73). The anomaly was highly visible in three geophysical layers (Figure 5.73a-d), and clearly evident in the ground (Figure 5.73e-f). Other wall anomalies were subtler than this when excavated, and included some adobe walls that are more characteristic of pueblo-style architecture.

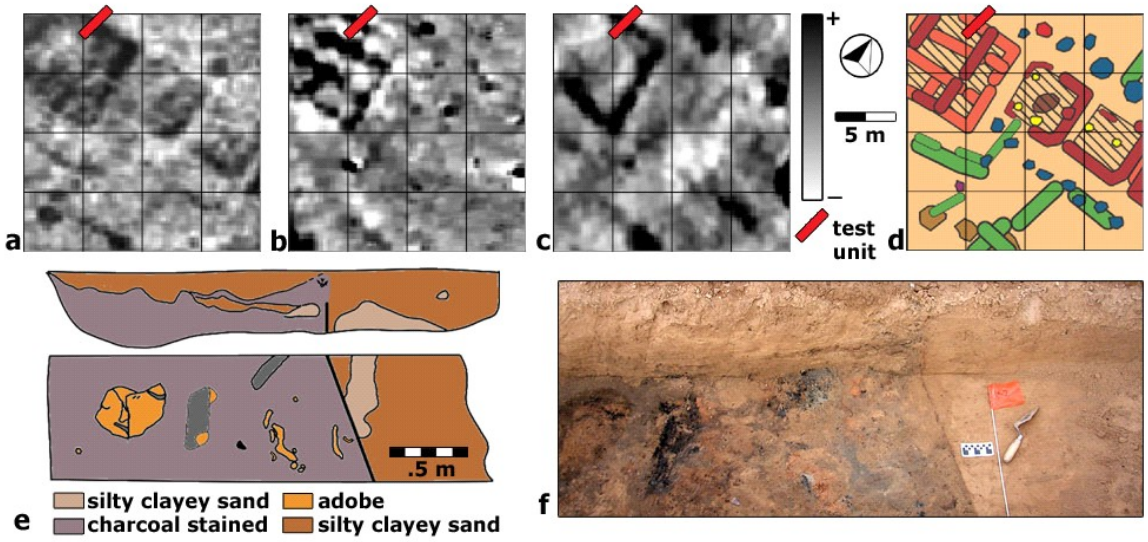


Figure 5.73. *Testing of a wall anomaly with mechanized trench M2. Small subareas surround the test unit are shown: a) GPR slice from .47-.63 m, b) magnetic gradiometry, c) magnetic susceptibility, and d) vector interpretation (for symbol key see Figure 5.68). e) Profile of the test unit's west wall and plan map of trench floor (adapted from Lukowski & Perez 2005). f) Photograph of central portion of the test unit's floor and wall.*

Interior room anomalies were relatively small compared with other anomalies, so many of them were sampled with cores rather than mechanized trenches or other large units. Six interior room anomalies were tested with cores and two with mechanized trenches, of which two were inconclusive, one provided no evidence of a feature, and five were verified. Although the use of cores made it difficult to describe the anomaly sources beyond the designation of "feature," the two mechanized trenches were enlightening. One revealed particularly dense cultural fill, and the other a hearth, which is illustrated in Figure 5.74. Unit M13 was placed entirely within the bounds of a probable house, with the anomaly of interest near the center. Ground-penetrating radar data of this area (Figure 5.74a) show a small structure with two small interior anomalies suggestive of post holes based on their size and spatial layout. Magnetic susceptibility and gradiometry data (Figure 5.74b,c) show an anomaly at the center of this structure, and this was the target for the test excavation. The excavation revealed what appears to be the floor of the structure with a concentration of charcoal and burned material near the center, confirming the target anomaly as a hearth. In addition, a probable post hole was discovered close to the center of the west wall. Although the GPR-indicated post hole appears to be outside the trench, it seems likely that the anomaly was actually excavated and is confirmed as a post hole. The positional discrepancy between the anomaly and the test unit is small and well within the expected range of error in the gridding system.

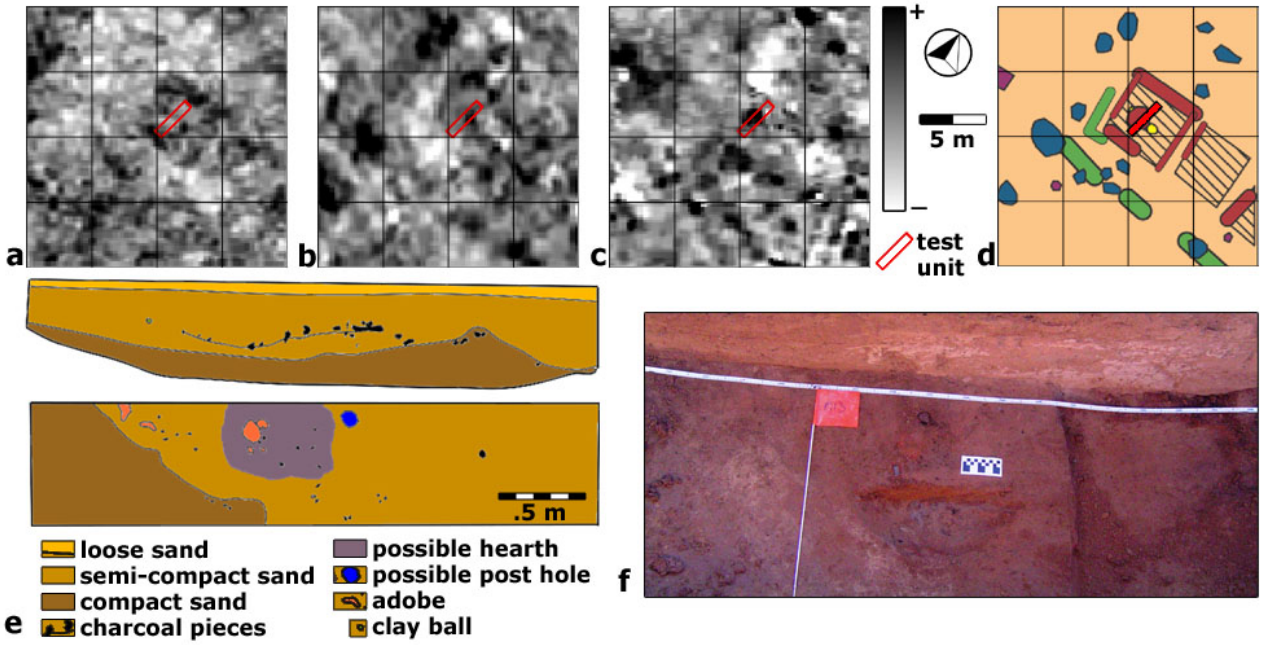


Figure 5.74. *Testing of an interior room feature anomaly with mechanized trench M13. Small subareas surround the test unit are shown: a) GPR slice from .31-.47 m, b) magnetic susceptibility, c) magnetic gradiometry, and d) vector interpretation (for symbol key see Figure 5.68). e) Profile of the test unit's east wall and plan map of trench floor (adapted from Lukowski & Perez 2005). f) Photograph of central portion of the test unit's floor, where the possible hearth has been bisected and sampled.*

Linear anomalies were distinct and prevalent in magnetic susceptibility and to a lesser degree in the GPR data, yet the majority of these could not be verified with test excavations. It should be noted that although these were considered linear, they were often composed of small, closely spaced amorphous anomalies. This is why it was suggested that these anomalies could be alignments of post holes or trash/storage pits. Of the nine linear anomalies tested, all with mechanized trenches, only two showed any evidence of a cultural feature. Not all of the remaining test units were "sterile," however, because three contained soil or stratigraphic anomalies that were thought to explain why an anomaly was created in the geophysical data. Given that these anomalies were parallel and at a rather constant distance from the house anomalies, it is actually more likely that they are directly related to the archaeology, if only as products of site erosion after abandonment. The best example of a verifiable linear feature was found in test unit M26, where a sediment layer enriched with ash and charcoal was discovered (Figure 5.75). The feature was clear in the east wall of the trench (Figure 5.75e), which is also where the anomaly is stronger (Figure 5.75a-c), but some evidence of the same staining was also evident in the west wall of the test unit.

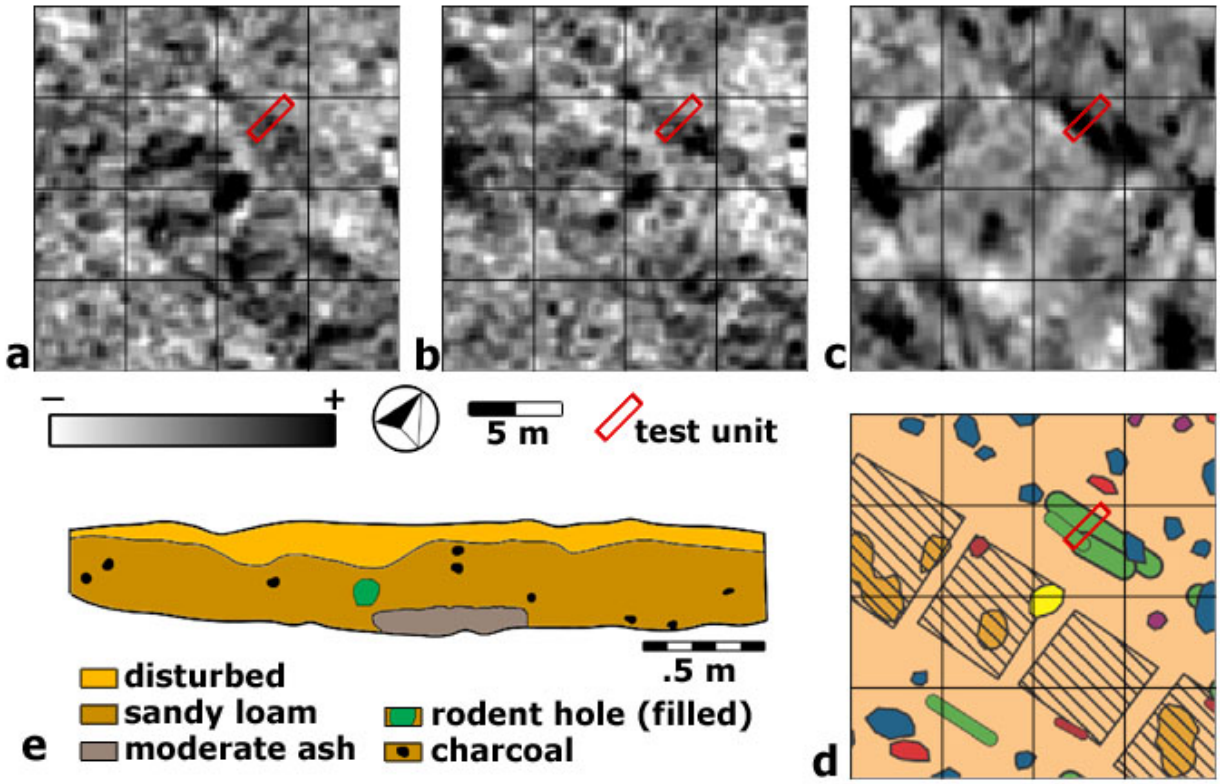


Figure 5.75. *Testing of linear anomaly with mechanized trench M26. Small subareas surround the test unit are shown: a) GPR slice from .31-.47 m, b) GPR slice from .47-.63 m, c) magnetic susceptibility, and d) vector interpretation (for symbol key see Figure 5.68). e) Profile of the test unit's east wall (adapted from Lukowski & Perez 2005).*

Small feature anomalies were met with mixed success. Of the eighteen tested, only eight were confirmed as cultural features, whereas nine were determined to be either non-cultural or non-existent as far as could be determined with excavations. One unit was excluded because it was not excavated deeply enough to fully test the anomaly. It is noteworthy that five of the units having no cultural features did contain something that explained why an anomaly was produced, including stratigraphic changes, rodent burrow, and root disturbances. Only two units, both mechanized trenches, found nothing that might explain a geophysical anomaly. Figure 5.76 illustrates two successful verifications of small features. Mechanized trench M25 was excavated to test one of the systematically distributed small features visible in GPR and magnetic susceptibility data (see Figure 5.76a-d), and the result showed part of a circular stain rich with cultural debris. This is interpreted as part of an outside activity area, patio, or midden deposit associated with the nearby room block. The discovery of this also sheds light on the nearby linear anomaly discusses above and illustrated in Figure 5.75, which fits into the context of a patio or similar activity and/or refuse area. One of the randomly distributed small feature anomalies is also illustrated (Figure 5.76g-j), in which the excavation of M19 revealed the remnant of a floor. This anomaly was only visible in magnetic susceptibility data (Figure 5.76g), but was clearly visible in the profile of the excavation (Figure 5.76i,j).

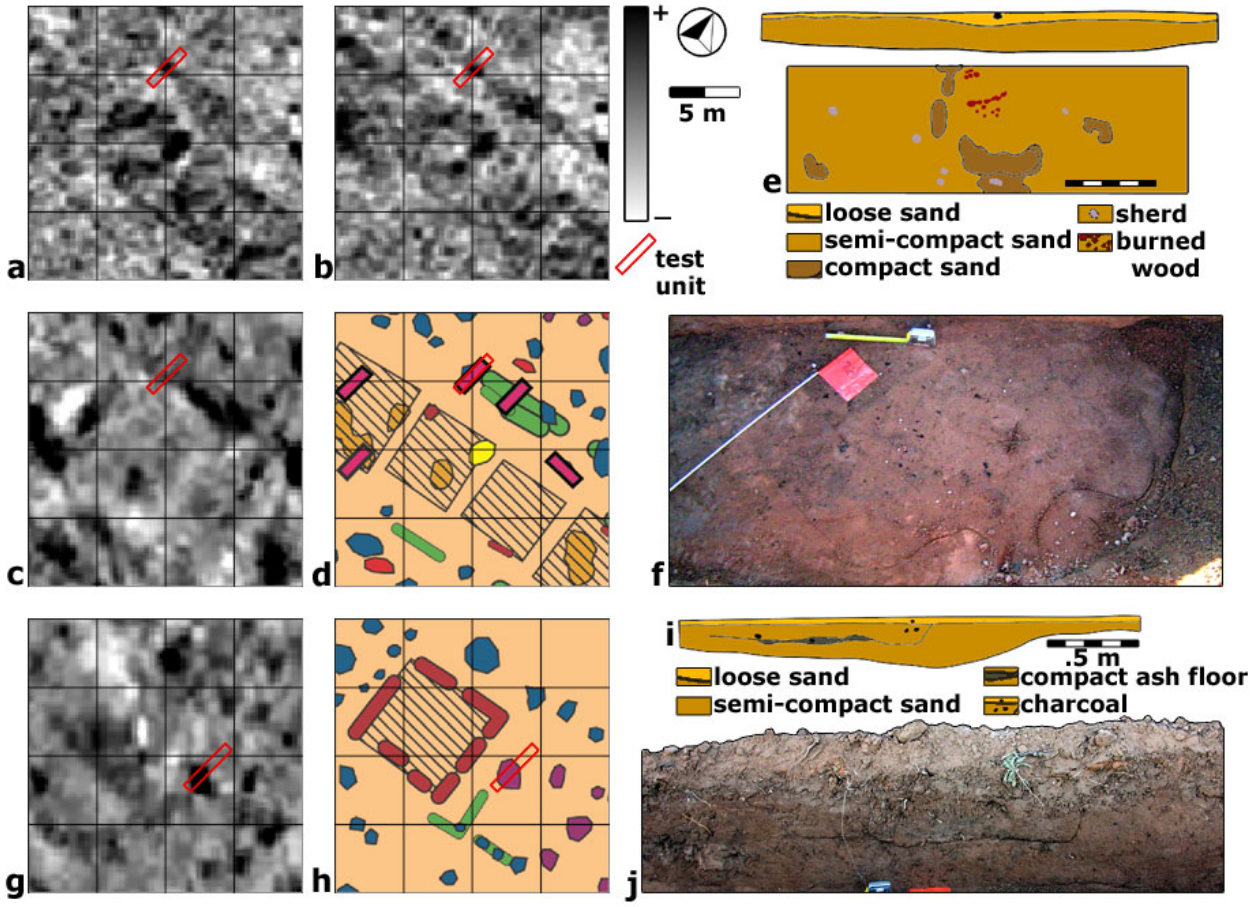


Figure 5.76. *Testing of two small feature anomaly with mechanized trenches M25 and M19. Small subareas surrounding M25 are: (a) GPR slice from .31-.47 m, (b) GPR slice from .47-.63 m, (c) magnetic susceptibility, and (d) vector interpretation (for symbol key see Figure 5.68). e) Profile of the M25's west wall and plan map of trench floor (adapted from Lukowski & Perez 2005). f) Photograph of central portion of the test unit's floor. Small subareas surrounding M19 are (g) magnetic susceptibility and (h) vector interpretation. i) Profile of M19's west wall (adapted from Lukowski & Perez 2005). j) Photograph of southern portion of M19's west wall.*

Table 5.36 shows the success of test excavations for each anomaly class, including percentage correct. Ninety-five percent confidence limits are also given for groupings of similar classes where sample sizes are somewhat larger. It is emphasized that statistical predictions must be viewed with caution when sample sizes are small, as discussed in Section 5.14. It is emphasized that these statistics would have been even higher had anomalies been hand picked for testing. Anomalies are commonly selected very carefully with only the best or most promising ones excavated. Instead, samples in this project were chosen randomly so that the statistical outcome would be more robust. Not surprisingly, anomalies that were predicted to be house floors, walls, and interior features were much more likely to be confirmed by test excavations than exterior lineations and small feature anomalies. The overall success rate for floors, walls, and interior room features (classes 1-9) was 85%, within 95% confidence limits ranging

between 69–93%. Lineations (classes 10-11) were least correct, with only 22% (95% confidence limits of 6-55%). The success of confirming small features fell between these two, with 47% correct and 95% confidence limits between 26-69%. The overall success rate (classes 1-14), excluding tests of the background, was 64% (95% confidence limits between 52-75%).

Table 5.36. *Summary of test excavation success at Pueblo Escondido.*

	pos	***neg***	***% correct***	***95% conf limits***	
01: Room block floor	4	1	80		
02: Room floor, burned	4	0	100		
03: Room block floor, inferred	2	0	100		
04: Outlying floor	5	0	100		
Floor total (1-4)	15	1	94	72	99
05: Room block wall, burned/magnetic	3	1	75		
06: Room block wall, robust	3	2	60		
07: Room block wall, subtle	2	0	100		
Wall total (5-7)	8	3	73	43	90
08: Interior room feature	2	0	100		
09: Interior room feature, burned	3	1	75		
Interior Room Feature total (8-9)	5	1	83		
Interior Feature total (1-9)	28	5	85	69	93
10: Lineation, robust	2	4	33		
11: Lineation, subtle	0	3	0		
Lineation total (10-11)	2	7	22	6	55
12: Small feature, systematic	5	3	63		
13: Small feature, random	1	4	20		
14: Small feature, burned	2	2	50		
Small feature total (12-14)	8	9	47	26	69
Extramural Feature total (10-14)	10	16	62	22	57
15: Background	11	0	100	74	100
Grand total (1-15)	49	21	70	58	79
Feature Total (1-14)	38	21	64	52	75

Abbreviations: “pos” stands for positive verification, indicating the number of predicted archaeological feature that were confirmed by excavation, and “neg” stands for negative confirmation, indicating the number of anomalies not confirmed; Confidence limits were calculated using an asymmetric approximation. Caution is advised when viewing statistics associated with low sample numbers.

Overall the predictions for Pueblo Escondido appear to be very successful, but some types of anomalies were not as reliable as others. To gain an appreciation for the advantage of using geophysics as compared to a random sampling approach often used in archaeology, the improvement over chance can be calculated. First, the probability of blindly discovering an archaeological feature must be estimated. A nearby and similar archaeological site, Firecracker Pueblo, provides a useful data set for comparison (O’Laughlin 2001). The site was excavated very thoroughly over a long period of time, and a map of the site, encompassing approximately .33 hectares, serves as the basis for

calculating the probability of finding an archaeological feature with a randomly placed .25 x .25 m excavation. Using this map, the chance of discovering an archaeological feature is only 29% (based on the area occupied by archaeological features divided by the total area exposed). It is important to note that this rate is rather high compared with circumstances at Pueblo Escondido. If the regression-based fusion (Figure 5.35a) is reclassified into feature presence and absence categories using the normal p=.5 as the cutoff, then the percentage of the area designated as having a feature is 16%. Using the Firecracker Pueblo map is more useful as a way to avoid biasing the results with information derived from geophysical data. Moreover, if the use of geophysics proves to be a significant improvement over chance, then the case is much stronger than if a more typical site with a lower feature density were used.

Using a binomial test for proportions comparing the theoretical probability based on Firecracker Pueblo (p=.29) and the overall probability of finding any of the features in categories 1-14 at Pueblo Escondido (p=.64), the use of geophysics represents a significant improvement over chance. The test statistic, a z-score of z=5.92, is well above the one-tailed critical value of 1.645 (n=59, α=.05). In fact, any proportion greater than p=.39 is a significant improvement over chance at α=.05.

The results of test excavations can also be used to evaluate the effectiveness of each geophysical data layer and of the types of excavation units utilized. Table 5.37 shows the percent correct rates and confidence intervals for each geophysical data layer that was used for the anomaly predictions. In this case, since many of the anomalies identified for testing were visible in more than one geophysical data layer, each has a relatively large sample size. Percentages of correct assignments for GPR (all slices together), magnetic gradiometry, and magnetic susceptibility are 77%, 79%, and 89%, respectively. Magnetic susceptibility was therefore the most reliable method. Magnetic gradiometry and GPR were almost equally effective, although the confidence interval for GPR is considerably narrower.

Table 5.37. *Summary of test excavation success by geophysical method at Pueblo Escondido.*

Data Layer	***pos***	***neg***	***% correct***	***95% conf limits***	
GPR 9-16 cm	8	3	73	43	90
GPR 15 – 31 cm	20	3	87	68	95
GPR 31 – 43 cm	19	5	79	60	91
GPR 43 – 67 cm	15	2	88	66	97
GPR (all slices)	27	8	77	61	88
Magnetic gradiometry	11	3	79	52	92
Magnetic susceptibility	24	3	89	72	96

See Table 5.36 for abbreviations. Caution is advised when viewing statistics associated with low sample numbers.

It has often been said that the reliability of an anomaly representing an actual subsurface archaeological feature increases when several geophysical methods show the same anomaly. This was true for the Pueblo Escondido data (Table 5.38). In the cumulative data, if only one type of data showed evidence for an anomaly, then the success rate was 57%. If the anomaly was visible in two data sets the percentage of

successes increased to 82%. Only four anomalies were indicated by all three geophysical data types, but all were successfully validated (i.e., 100%).

Table 5.38. *Summary of test excavation success according to the number of different geophysical indications. Three indications means that the anomaly was visible in magnetic susceptibility, magnetic gradiometry, and at least one GPR slice. Two indications means two of the three geophysical data types showed the anomaly, and one indication means only one type showed the anomaly.*

	pos	***neg***	***% correct***	***95% conf limits***	
1 indication	25	19	57	42	70
2 indications	9	2	82	52	95
3 indications	4	0	100	51	100

See Table 5.36 for abbreviations. Caution is advised when viewing statistics associated with low sample numbers.

The excavation results also allow a comparison of the effectiveness of different types of excavation units (Table 5.39). Units exposing areas, which include the mechanized trenches plus the two shovel scrapes, were only slightly more effective at 76% (with a 95% confidence interval between 61-86%) than core holes at 71% (with 95% confidence interval between 45-88%). It is important to remember, however, that interpretation of core holes was done after the other excavation units had been completed and the knowledge gained from several days of excavation greatly aided the ability to interpret the them. Core holes alone would probably be much less effective for most types of archaeological features.

Table 5.39. *Summary of test excavation success by type of excavation used at Pueblo Escondido. Data for testing the background (15) were excluded. Aerial units include mechanized trenches and shovel scrapes.*

	pos	***neg***	***p***	***95% conf limits***	
Aerial Units	34	11	0.76	0.61	0.86
Cores	10	4	0.71	0.45	0.88

See Table 5.36 for abbreviations. Caution is advised when viewing statistics associated with low sample numbers.

All of the above statistics were calculated excluding the eleven core holes that tested the "background" class, which includes all areas devoid of anomalies (approximately 90% of the survey area). Not one of these core holes found adequate evidence for an archaeological feature. In fact, ten of them were "sterile" and one contained only a few small bits of charcoal. Although the background test was very limited in extent, the outcome supports the utility of geophysical data to locate not only the presence of archaeological features, but to accurately indicate areas of very low feature density or altogether absence.

Discussion. The Pueblo Escondido test excavations shed light on the perceived ability to validate geophysical anomalies with excavations and on what forms of evidence are valid for archaeological interpretation. While a great many predicted features were verified as cultural in nature by excavation, some turned out to be minor changes in stratigraphy, rodent holes, and root disturbances. Most importantly, some excavations revealed no indication of anything that might explain the cause of an anomaly. While it is

true that some geophysical anomalies are the product of instrument or user error, or data processing artifacts, many are clearly not accidental. The Pueblo Escondido data provide several instances where clearly cultural anomalies were difficult or impossible to verify with excavations. The robust lineation tested by unit M26 is a case in point (Figure 5.75). TRC reported that they found "no evidence of the lineation" in the excavation unit, and only scattered pieces of charcoal (Lukowski & Perez 2005:63). Yet, when the same excavation unit was examined by SERDP participants, a discrete lens of ash-stained sediment was found in the eastern trench wall that was determined to be the cause of the anomaly. This is one of the many cases in which TRC did not find evidence of the targeted feature, but SERDP members did. It is assumed that this is in no way an indication that TRC was unscrupulous, but instead that many of the anomaly sources were atypical and would have been overlooked by most archaeologists. SERDP team participants made a strong effort to identify all disturbances in the ground that might be the cause of the anomalies.

Another similar example was found in test unit M23, where GPR data show a very distinct rectangular outline that was interpreted as the walls of a house (Figure 5.77a-d). Inside this rectangle is a large magnetic susceptibility anomaly, suggesting cultural fill (Figure 5.77e). Yet TRC did not find any evidence of cultural features, and even expressed concern for this fact because the anomaly is so clear and distinct (Lukowski & Perez 2005:63). They suppose that perhaps the unit was incorrectly located. This is unlikely however, and the SERDP team came to different conclusions. A basin-shaped area with laminations of fine sand was discovered in the trench wall. Since no cultural materials were found within these layers, they are probably the result of natural infilling of a structure that was rather empty when it was abandoned. This anomaly was documented by TRC as unverified, but it is concluded here that this was an accurate finding for the predicted anomaly. Without the benefit of the geophysical data, the excavation findings might not be very strong evidence for a house, but the clear geophysical indications strongly support the final interpretation—natural phenomena cannot create an anomaly of this form.

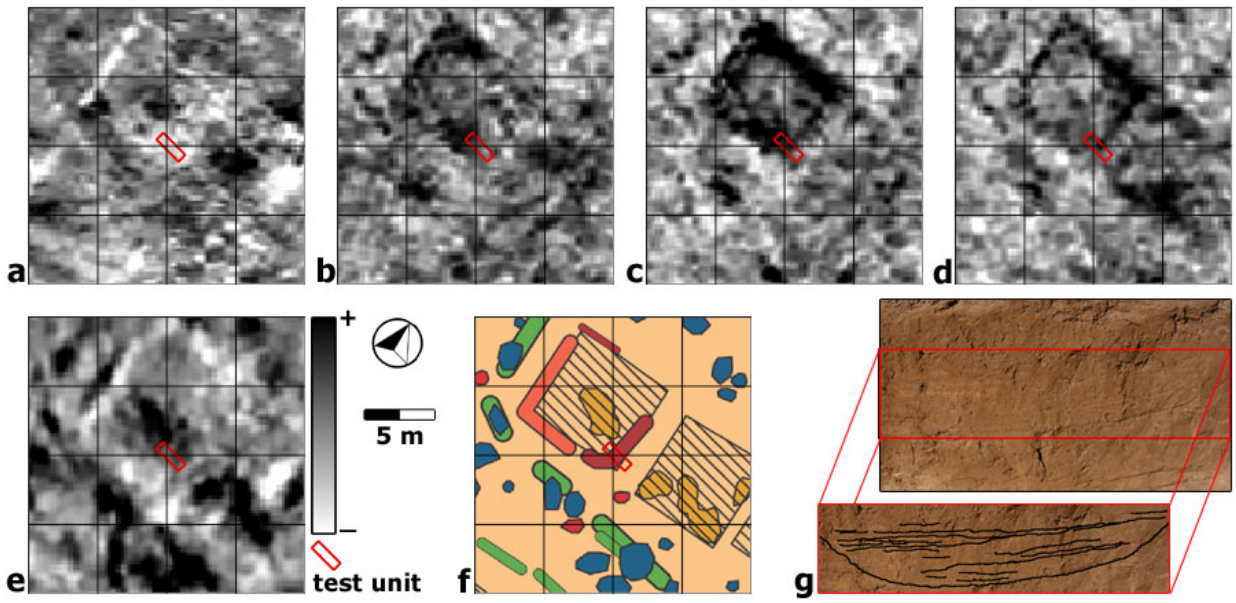


Figure 5.77. *An example of distinct house anomaly that was nearly invisible in the test excavation. Small sub-areas surround the test unit are shown: a) GPR slice from .09-.16 m, b) GPR slice from .15-.31 m, c) GPR slice from .31-.47 m, d) GPR slice from .47-.63 m, e) magnetic susceptibility, and f) vector interpretation (for symbol key see Figure 5.68). g) Photograph of the test unit's north wall, showing laminations of fine sand in a basin-shaped depression.*

Yet another example is seen in excavation unit M28 (Figure 5.78), where GPR data very clearly show the square plan of a buried structure. TRC did not find any archaeological evidence aside from some scattered charcoal, one flake and one sherd (Lukowski & Perez 2005:64). The SERDP team did, however, notice a dry, compact sediment layer that was interpreted as a possible house floor. Still, this evidence was not very strong because the compact surface had no lateral boundaries, so it was concluded by both groups that the anomaly could not be confirmed. (There is the possibility that excavations did not go deep enough, but this would only be possible if GPR depth calculations are greatly in error.) It was agreed that there was nothing visible in the excavation to suggest the presence of a cultural feature. Yet the anomaly is so clear and most certainly represents a buried structure. The fact that it could not be detected with excavations underscores the notion that some types of archaeological features cannot be verified with traditional excavation methods based on sight and touch.

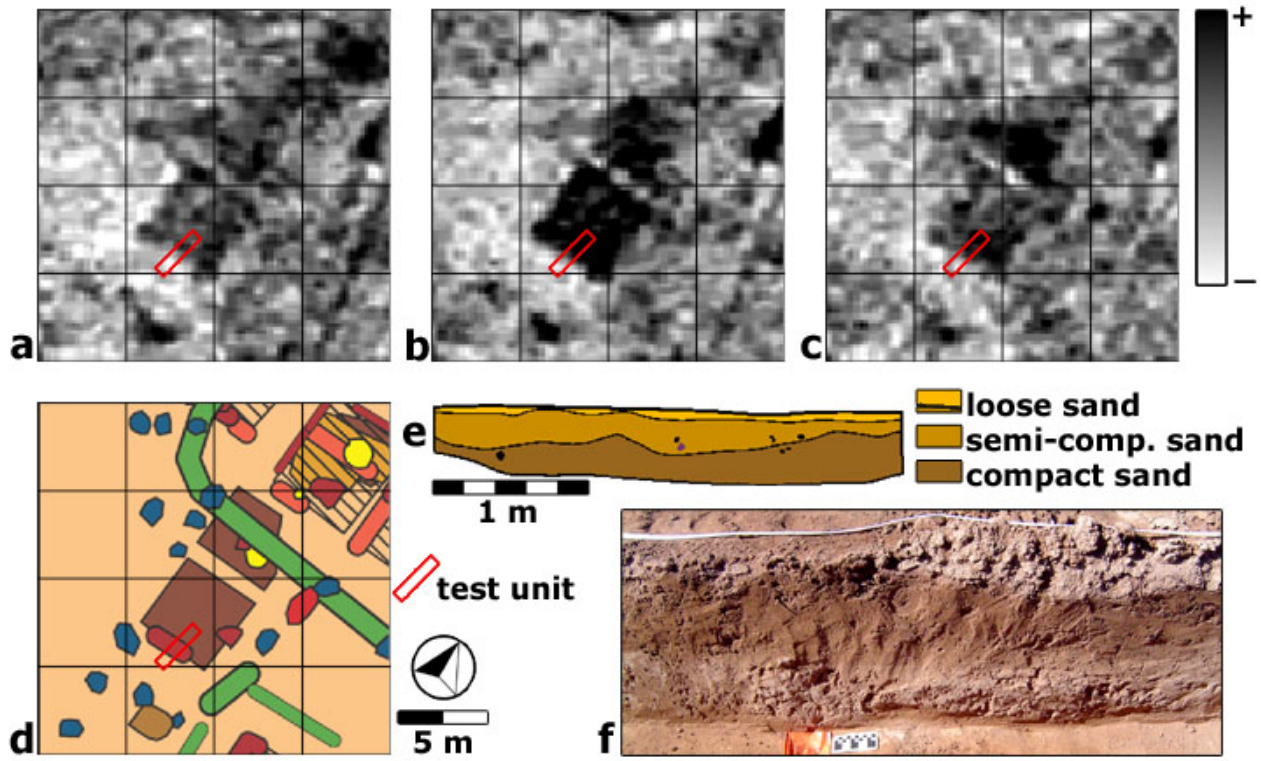


Figure 5.78. *An example of a distinct house anomaly for which no evidence was found by excavation (unit M28). Small subareas surround the test unit are: a) GPR slice from .15-.31 m, b) GPR slice from .31-.47 m, c) GPR slice from .47-.63 m, and d) vector interpretation (for symbol key see Figure 5.68). e) Profile showing stratigraphy (adapted from Lukowski & Perez 2005). f) Photograph of the test unit's west wall.*

Findings in test units M26 and M23 are not the only examples of contradictory evaluations between TRC and the SERDP team. In fact, these kinds of situations are the main reason for the large differences in overall success rates reported by TRC compared to those reported in this chapter. The discrepancies are very enlightening, because they show that geophysical data can show very "real" archaeological features that are otherwise not readily detectable by traditional excavations. They also strongly suggest that archaeologists should pay more attention to subtle sedimentary changes and should view them as at least potentially cultural. It is suggested here that the use of traditional archaeological excavations may be an insufficient way of validating geophysical anomalies in some instances. A new form of field protocol may be necessary. Vast improvements could be made through use of hand-held geophysical sensors, soil sample analyses, and the expertise of a geoarchaeologist.

5.16.3. Field Investigations at Silver Bluff

Michael L. Hargrave, Construction Engineers Research Laboratories (CERL)

Ground truthing investigations at Silver Bluff (the George Galphin site) were conducted by the Savannah River Archaeological Research Program (SRARP), an

affiliate of the University of South Carolina. The investigations were directed by Ms. Tammy Forehand Herron with assistance from Mr. Robert Moon, a SRARP crew of three, and four volunteers. Ms. Herron was uniquely qualified to direct the study because she had conducted ongoing research at Silver Bluff for several years, and was thus very familiar with the site's soils and feature types. Fieldwork occurred from March 21-25, 2005. The SERDP team was represented by Kvamme and Hargrave, with assistance from volunteer JoAnn Kvamme, an experienced archaeologist and archaeo-geophysicist.

The geophysical survey area at Silver Bluff included approximately twenty-four and one-half 20 x 20 m blocks that covered a total area of 9,800 m^2. The excavations consisted of 20 shovel test pits, two hand-excavated units measuring .5 by 1 m, and 50 backhoe trenches (Figure 5.53). Excavation and documentation protocols were essentially the same as those described for the other sites. All of the shovel test pits were targeted on magnetic anomalies that were assumed to represent either historic artifacts or recent ferrous metal objects, particularly wire pin flags. The backhoe trenches investigated a random sample of the other anomalies as defined by the sampling program (Table 5.22; Figure 5.79).

Figure 5.79. *Excavation of a backhoe trench at Silver Bluff, March 2005. Dense undergrowth was burned shortly before fieldwork began.*

At Silver Bluff, with less understanding of the nature and variability of the subsurface archaeology, the geophysical anomalies were assigned to categories that were descriptive in nature (Section 5.14.4). Aside from the metal-based point anomalies, most of the defined types were linear although some area anomalies were also selected for investigation (Table 5.22). In some cases the SERDP team provided additional inferences about *possible* archaeological feature types that could be associated with a specific class of anomalies. For example, some anomalies were identified as possible floors or interior spaces, storage features, lanes, and so forth. The many subtle linear GPR anomalies could represent a low wall of brick or stone, a low berm of packed earth, a compacted trail or path, a narrow paved walkway, a line of contiguous post holes, a line of somewhat

separated post holes, or a narrow slit trench or builder's trench within which wooden constructions once were placed (Section 5.12.6).

5.16.3.1. Evaluation of predictions by archaeological team
Michael L. Hargrave, CERL

During fieldwork, archaeological features were identified in many of the excavation units, but in most cases it was difficult to determine whether they were causally associated with the targeted anomalies. Some of this uncertainty was due to the limited amount of excavation. A great deal of time was spent using trowels and shovels to better define features, but the budget did not permit them to be excavated (Figure 5.79). Some stains were described as "possible features" (Herron and Moon 2005), whereas others were too problematic for even that designation. Even if it had been possible to completely excavate a sample of Silver Bluff's features by hand, however, this would probably not have resolved the issue of whether they explained the presence of particular geophysical anomalies. For example, posts were the most common type of feature documented at the site, occurring in 18 of the excavation units (Table 5.40). Posts are often associated with walls, and walls are a potential explanation for linear geophysical anomalies. It is conceivable that the only archaeological manifestation of a wall might be those few posts that happen to extend the deepest below surface, or that include relatively dark fill. Archaeological sites frequently also include many isolated posts. Some of these are often assumed to be associated with extramural facilities (e.g., drying racks, sun screens) whereas others may, in fact, have non-cultural origins (including tap roots, rodent burrows, etc.). Assuming that the presence of one or two posts identified within the narrow confines of a backhoe trench indicates that a wall was present thus requires a considerable inferential leap. Similarly, midden (soil enriched by the introduction of culturally-derived organic materials and artifacts) was identified in a number of the trenches at Silver Bluff, and a linear band of midden might explain a linear anomaly. Unfortunately, time did not permit most trenches to be expanded sufficiently to determine whether a linear midden deposit was present. While interpretive problems of this type complicate archaeological investigations at many sites, they were particularly troublesome at Silver Bluff because of the abundance of linear anomalies.

Table 5.40. *Cross tabulation of anomaly categories and feature types at Silver Bluff.*

Feature Type Present		Iron or Steel art	Post (s)	Post & Midden	Post & Pit	Midden	Pit	Tree	Post & Tree	Shovel Test	Shovel Test & Post	Linear Feat	Linear Feat & Post	Lane?	Other Feat	No Feat	TOTAL
Anomaly Category:																	
Historic iron or recent steel artifact	n	20															20
	%	100.0%															100.0%
Linear negative resistance	n		1	1		1										3	6
	%		16.7%	16.7%		16.7%										50.0%	100.0%
Linear positive resistance	n		1					1				1		1	1	5	10
	%		10.0%					10.0%				10.0%		10.0%	10.0%	50.0%	100.0%
Linear subtle GPR	n		4	1	1		1		1	1	1			1	2	5	18
	%		22.2%	5.6%	5.6%		5.6%		5.6%	5.6%	5.6%			5.6%	11.1%	27.8%	100.0%
Linear subtle negative GPR	n															1	1
	%															100.0%	100.0%
Linear robust GPR	n		1	1	1	1										1	5
	%		20.0%	20.0%	20.0%	20.0%										20.0%	100.0%
Linear subtle magnetic	n													1		1	2
	%													50.0%		50.0%	100.0%
Areal negative resistance (floor or inte	n						1								1	1	3
	%						33.3%								33.3%	33.3%	100.0%
Areal negative resistance (uncertain ty	n															1	1
	%															100.0%	100.0%
Areal positive resistance (floor)	n												1				1
	%												100.0%				100.0%
Areal GPR (floor or storage area)	n				1			1									2
	%				50.0%			50.0%									100.0%
Magnetic point non-dipole in alignmen	n		1	1													2
	%		50.0%	50.0%													100.0%
Lightening strike	n															1	1
	%															100.0%	100.0%
Broad deep zone robust GPR (lane?,	n					1											1
	%					100.0%											100.0%
TOTAL	n	20	8	4	3	3	2	2	1	1	1	1	1	3	4	19	73
	%	41.1%	43.5%	1.4%	1.0%	1.0%	0.7%	0.7%	0.3%	0.3%	0.3%	0.3%	0.3%	1.0%	1.4%	6.5%	100.0%

Notes: 2 of the 20 iron or steel artifacts are described as metal sheeting and "tin" can.

As a consequence of these interpretive difficulties, Herron and Moon (2005) did not attempt to quantify the rates of accuracy for anomaly categories in the manner that was done at Army City and Pueblo Escondido. Table 5.40 was compiled for this report using the narrative information presented by Herron and Moon (2005). The table provides an opportunity to examine the types of features that were identified in units excavated to investigate particular anomalies. One way to evaluate the validity of the anomaly categories is to consider how frequently the excavation units yielded no evidence for archaeological features. No features were located in 19 (35.9%) of the 53 units (the 20 shovel test excavations are excluded, as they were all targeted on anomalies interpreted as ferrous metal artifacts rather than soil features). The general prediction that a cultural feature was present was thus verified in 64.1% of the cases. Note, however, that this does not mean that 64.1% of the predictions based on the geophysical data were verified. In many cases it remains unclear if the documented features explain the presence of the targeted anomalies.

Descriptive anomaly categories would be very useful to archaeologists if they were found to always (or nearly always) be associated with a particular type of feature. This is the case for the category “Historic iron or recent steel artifact.” Iron artifacts such as nails or recent steel objects like pin flags were found in 18 of the 20 shovel tests. At Silver Bluff it was clearly possible to reliably map the distribution of ferrous artifacts. This capability could have made greater contributions to an understanding of activities conducted at the site in the 18th century if the ferrous metal distributions were not complicated by clutter associated with pin flags (see Section 3.3.4.3).

Three anomaly categories (linear negative resistance, linear positive resistance, and linear subtle magnetic) were not associated with features in 50% of the units. In most of these 8 units it was noted that the plow zone was relatively thin or there was a rise in the subsoil (Figure 5.80). Localized variation in the relative thickness of the soil strata

may thus account for these linear anomalies. This variation may be associated with plow furrows, but this remains speculative.

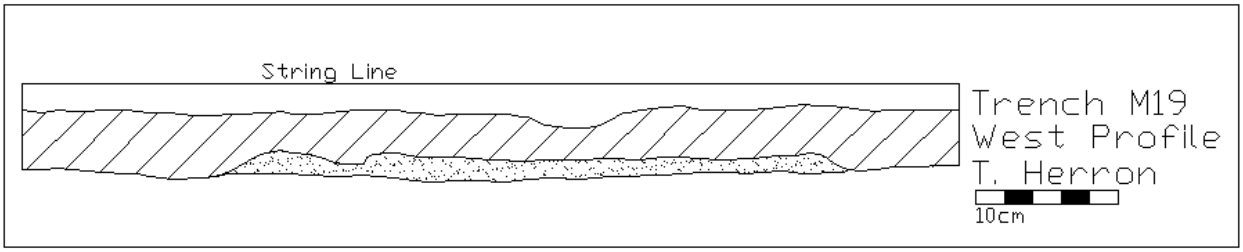


Figure 5.80. *Profile map of Mechanized Trench 19. Note how the lower soil stratum is nearer the surface in the central portion of the profile.*

Unfortunately, sample size limits one's ability to evaluate nearly all of the other anomaly categories. All of the cells in Table 5.40 that contain percentages of 50% or greater (other than the 'iron/steel' and 'no feature' categories) have samples sizes of only 1-2. Larger sample sizes can be achieved by combining categories. For example, the three types of linear GPR anomalies were investigated by a total of 24 excavation units. Eleven of these (45.8%) included at least one post hole. Post holes occur in only 18 (34%) of the 53 non-shovel test units, so their more common association with linear GPR anomalies may be meaningful.

In summary, it was particularly difficult at Silver Bluff to determine whether features such as post holes, pits, and midden should be viewed as the cause of geophysical anomalies (Figure 5.81a). Many of the anomalies were clearly linear, whereas most of the features were not. The ground truthing excavations resulted in the identification of "plow sole" deposits that had not been previously documented at the site. These reddish-colored sediments are believed to form at the base of plow furrows as a result of repeated plowing to the same depth. This indication of intensive plowing may have implications for the abundance of linear resistance anomalies that do not appear to denote architectural features (Figure 5.81b).

The results of ground truthing at Silver Bluff were surprisingly disappointing. While it was immediately apparent when first visiting the site that the extensive use of metal pin flags would diminish the interpretability of the magnetic data, the GPR and resistance surveys identified many anomalies of likely anthropogenic origin. The geophysical maps suggested that the core area of the trading post had been very heavily modified by human occupation. It was thought that this might be due to the presence of a palisade that strongly constrained the use of space. Surprisingly, the backhoe trenches provided little evidence for such intensive occupation. Historical records indicate that the area was under continuous cultivation from the mid-18th to the later 20th century, even though the soil is not well suited for such use. This use may explain the preponderance of linear anomalies, although the exact mechanisms involved are not known. Admittedly, the linear anomalies do not appear to represent individual plow furrows in size or orientation.

The SRARP excavators, who have been investigating the site for a number of years, noted that the geophysical data were useful in verifying the presence and extent of features that had been minimally defined in previous excavations, and by identifying a number of previously unknown features (Herron and Moon 2005).

a b

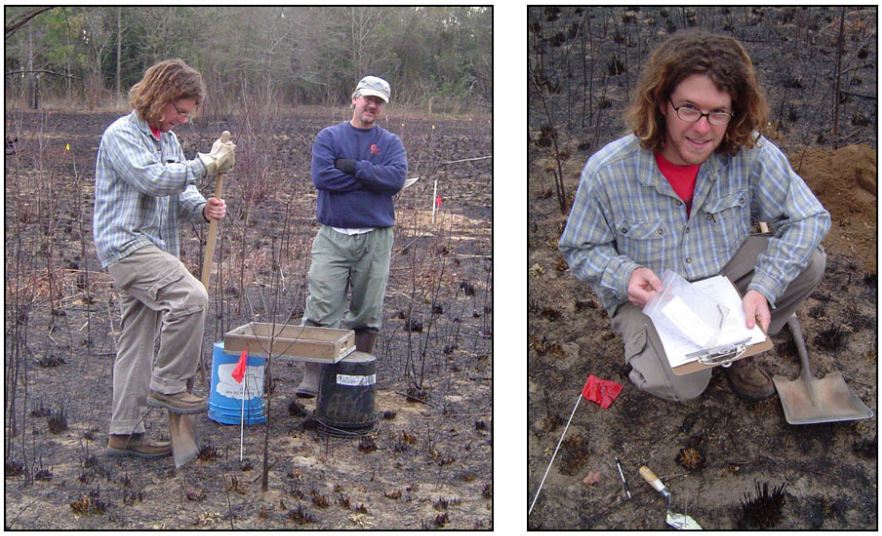
Figure 5.81. *Features revealed during Silver Bluff excavations. a) A large brick-filled post hole that may have functioned to buttress the palisade appears in the left foreground of Trench 27 at Silver Bluff. b) Prominent plow scars in Trench 17 at Silver Bluff.*

5.16.3.2. Evaluation of predictions by SERDP team
Kenneth L. Kvamme, University of Arkansas

As at the other sites the SERDP team performed a largely independent evaluation of the archaeological evidence for or against the assigned anomaly types. As observed above, unlike the previous sites, linking features exposed by excavation with the defined anomalies and arguing causal relationships proved to be very difficult at Silver Bluff. Many features in the soil occurred as simple amorphous stains that were difficult to define as specific types of archaeological feature, putting much of the following evaluation on shaky ground. Furthermore, it was very difficult to learn much more about this site through the excavation process.

Although most of the anomaly classes were very generic (Table 5.22), one named a specific class of cultural feature: iron or steel artifacts. In a historic site of this nature one would expect numerous iron artifacts across the site use area. The magnetic gradiometry survey certainly supports this idea with countless dipolar anomalies (Figure 4.45), but in Section 4.6.3 and elsewhere it was complained that numerous fallen pin flags and pin flag parts (from mowing) also litter the site. Although hundreds of pin flags were harvested prior to the geophysical surveys, the question remains of how much of the magnetic evidence points to historic sources as opposed to the recent pin flag influence? In the sampling design (Section 5.14.4) 189 major dipolar magnetic anomalies were defined via a threshold, and a simple random sample of 20 were selected for field testing by simple STP because these targets are easily located by metal detector (Figure 5.82). Surprisingly, many of these anomalies do, indeed, represent historic artifacts (all

recovered were nails), although an equal number signify pin flags or modern steel rubbish (Table 5.41). Based on this evidence, with *half* the samples showing historic artifacts, we believe the magnetic depiction of Figure 4.45 may give a reasonable portrayal of the distribution of historic ferrous iron in the site. Ninety-five percent confidence limits for the percentage of dipolar anomalies representing historic artifacts based on *n*=20 trials range between 28-72%.



a b

Figure 5.82. *Archaeological testing of the "iron or steel" anomaly category. a) Beginning excavation of a STP at the flagged locus of the pin flag; note screen for sifting through soil. b) Recovery of an historic nail, the source of the anomaly.*

Table 5.41. *Distribution of iron and steel artifacts in STP.*

Pin flags only	Other recent steel artifacts	Pin flags + historic nails	Historic nails only	Total
7	3	4	6	20

Another magnetic anomaly type that appeared to be detectable in the ground was a collection of small point anomalies in linear arrangement (Figure 5.83). Exposure by excavation revealed a series of small soil stains dark in color and roughly circular in shape (Figure 5.83b,c). The matrix contains a few small artifacts and charcoal and the features varied in thickness between 10-19 cm. Given the sea of ferrous metal-caused anomalies in the gradiometry data it is noteworthy that any soil features are discernable. As noted in Section 5.15, magnetic susceptibility data were collected in many archaeological features using a hand-held KT-9 Kappameter. To record data, one merely places the sensor head on the floor to measure magnetic susceptibility (MS) in SI units (Dalan 2006). Because this feature type is magnetic, MS measurements are relevant here. In trench M11, MS measured $.2x10^{-3}$ SI within the right-most features in Figure 5.83c and $.1x10^{-3}$ SI in the normal background (orange sandy soil). In trench 23 (Figure 5.83b), MS was recorded at $.26x10^{-3}$ SI with the background at $.14x10^{-3}$ SI. These data help explain the magnetic contrast that created the anomalies recorded by the magnetic gradiometry survey (Figure 5.83a). What the stains represent is open to question; their linear arrangement strongly suggests they are anthropogenic and they appear to coincide with a GPR revealed lineation (Figure 5.83a). A final magnetic anomaly type was

continuous and linear, but very subtle, with only 13 defined in the survey area and two tested in the site (Table 5.22). One excavation produced a dark cultural stain (MS measured .26x10^{-3} SI compared to a background of .15x10^{-3} SI) while the other was more ambiguous, ill-defined, and less certainly an archaeological feature (MS measured .26 x10^{-3} to .38x10^{-3} SI compared to a background of .15x10^{-3} SI).

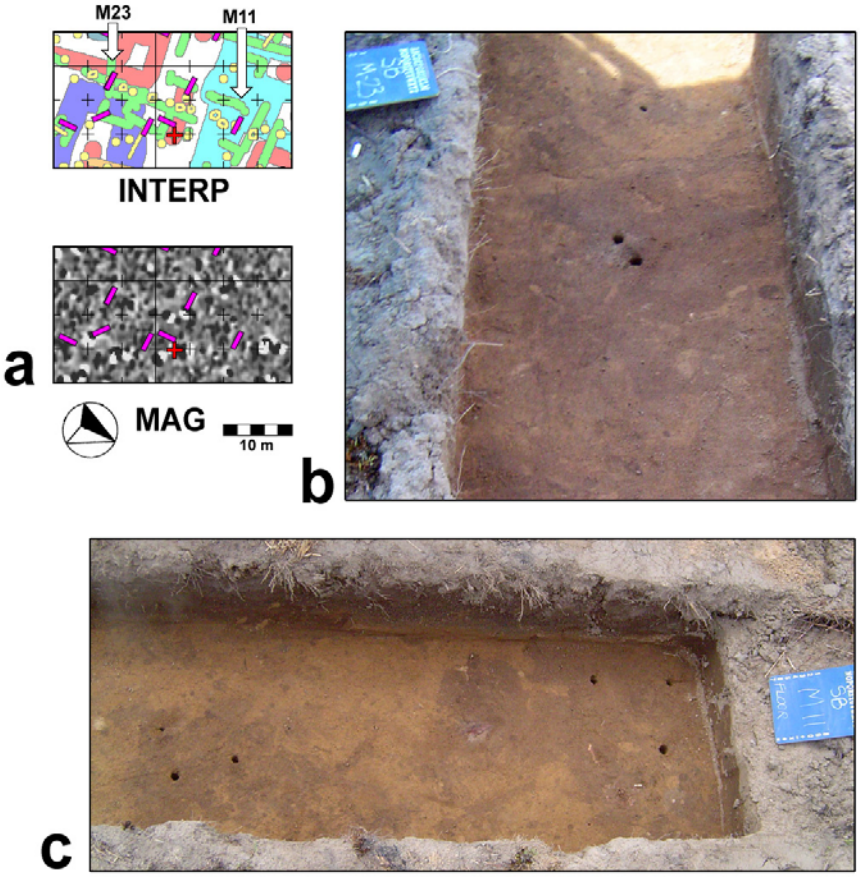


Figure 5.83. *Evidence for the anomaly type "magnetic point anomalies in linear arrangement." a) The magnetic and interpreted anomalies (in yellow). b) Anomaly source exposed in trench M23. c) Anomaly sources exposed in trench M11. The small holes in the excavation floors are from Oakfield cores used to explore feature depth. KEY: INTERP=interpretations; MAG=magnetic gradiometry.*

Many of the defined anomalies at Silver Bluff were classified as "area" features—indicated either by resistivity or GPR survey. It was believed that those portrayed by GPR might represent significant floors of brick, gravel, or packed earth, for example. Most of the area anomalies seen in resistivity were "negative," with atypically low measurements. Such anomalies point to more conductive soil that might occur within a sunken house floor or a midden, for example, where culturally derived sediments and greater ground moisture might accrue. Several of these area anomalies and their corresponding archaeological evidence are presented in (Figure 5.84). They are all revealed as dark stains of culturally enriched soil 35-45 cm thick. All three are revealed as low resistivity anomalies. Compared to the surrounding orange sandy soil, this material must be highly conductive, lowering resistivity. The soil stain in trench M30 (Figure 5.84c) was also expressed as a strong and rare GPR area anomaly (Figure 5.84a). Its color appears different, and a coarser and more compact texture was recorded.

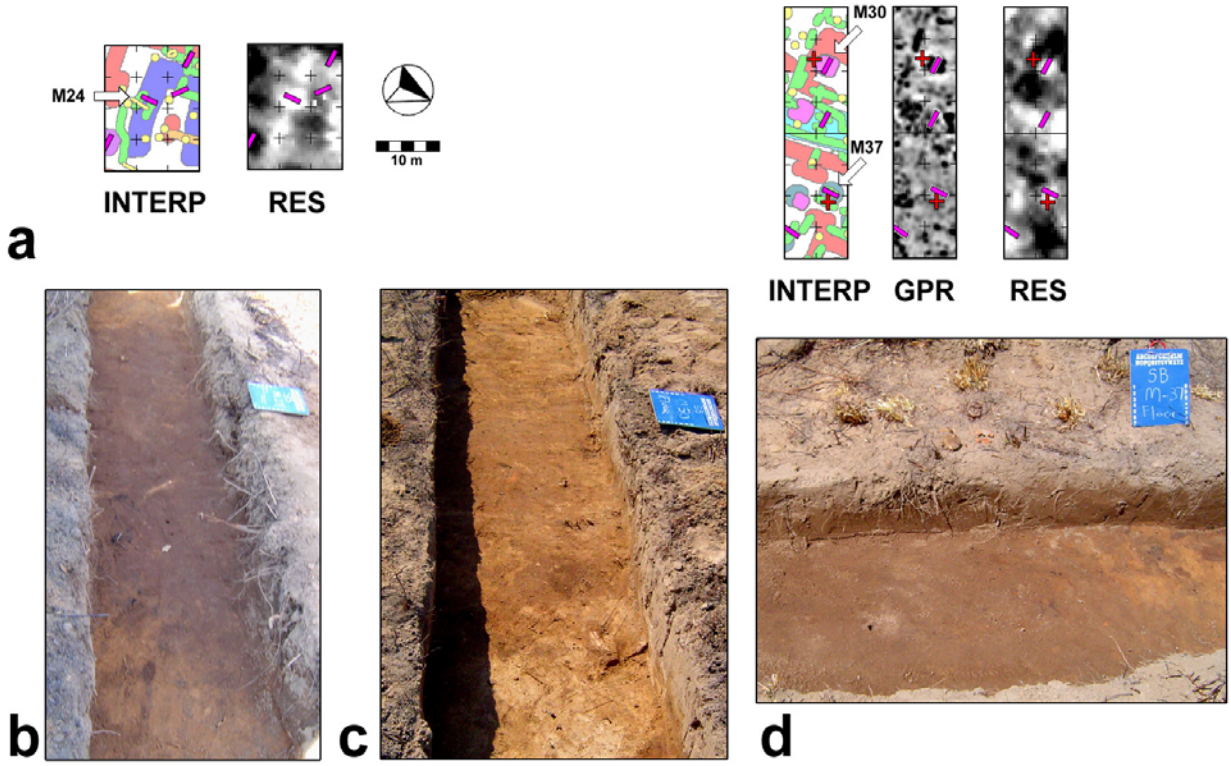


Figure 5.84. *Area anomalies and corresponding evidence at Silver Bluff. a) Geophysical evidence and interpretive maps. b) Trench M24 showing pronounced stain. c) Trench M30 showing pronounced stain. d) Trench M37 showing pronounced stain. KEY: INTERP=interpretations; RES=resistivity; GPR=ground-penetrating radar.*

Most of the anomalies at Silver Bluff were linear in form and were therefore interpreted to represent linear archaeological features of various types (see above). Lineations revealed by resistivity were particularly broad (often up to 4 m in width; Figure 4.45) and were defined as "positive" or "negative" (Section 5.12.6). It was speculated that many of the latter might represent lanes or small tracks running between various structures or work areas that once populated the site. The former were thought to represent adjacent (and resistant) low berms of soil. The archaeological sources of these anomalies were extremely difficult to identify and few of these anomalies could be corroborated with field evidence (see below). Possible sources of positive resistivity anomalies are illustrated in Figure 5.85. A circular region of bricks may explain a positive area anomaly in trench M27 (Figure 5.85b). In Figure 5.85c, the only possible explanation for high resistivity in this unit is the thinned plow zone or the rising and more resistant sandy subsoil. Several examples of this phenomenon were located. Given the extensive farming that is known to have occurred on this site (amply illustrated by plow scarring in many features), it is quite possible that any raised berms of soil that may have paralleled hypothetical lanes, farming plots, or work areas have been effectively leveled leaving geophysics as the only means by which such structure may now be easily detected (see below for further discussion of this idea).

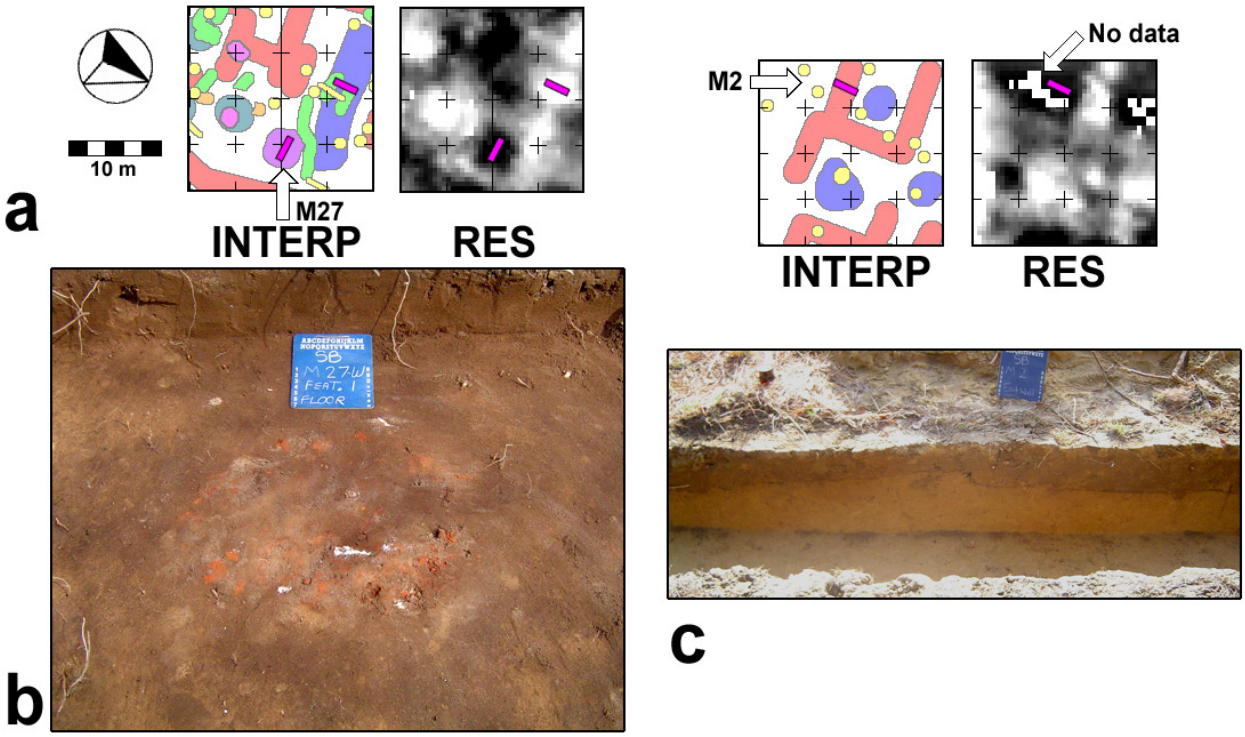


Figure 5.85. *Positive resistivity anomalies and possible archaeological sources. a) Geophysical evidence and interpretive maps. b) Collection of resistant bricks in trench M27 surrounding a post hole. c) Rising resistant subsoil in trench M2. KEY: INTERP=interpretations; RES=resistivity.*

The most prevalent anomaly types (after isolated magnetic dipoles) were linear and defined by GPR as "robust" and "subtle." Several, from both groups, are illustrated in Figure 5.86. Portions of two robust linear GPR anomalies are illustrated in (Figure 5.86a,b). Their extreme width and complex compound deposits are characteristic. Both of those illustrated may form parts of the outer palisade of Galphin's Trading Post. Subtle linear GPR anomalies and their archaeological sources are illustrated in (Figure 5.86c-e). These anomalies are typically harder to see and show less width, and this is generally typical of their soil expressions. All consist of relatively simple soil stains of varying contrast.

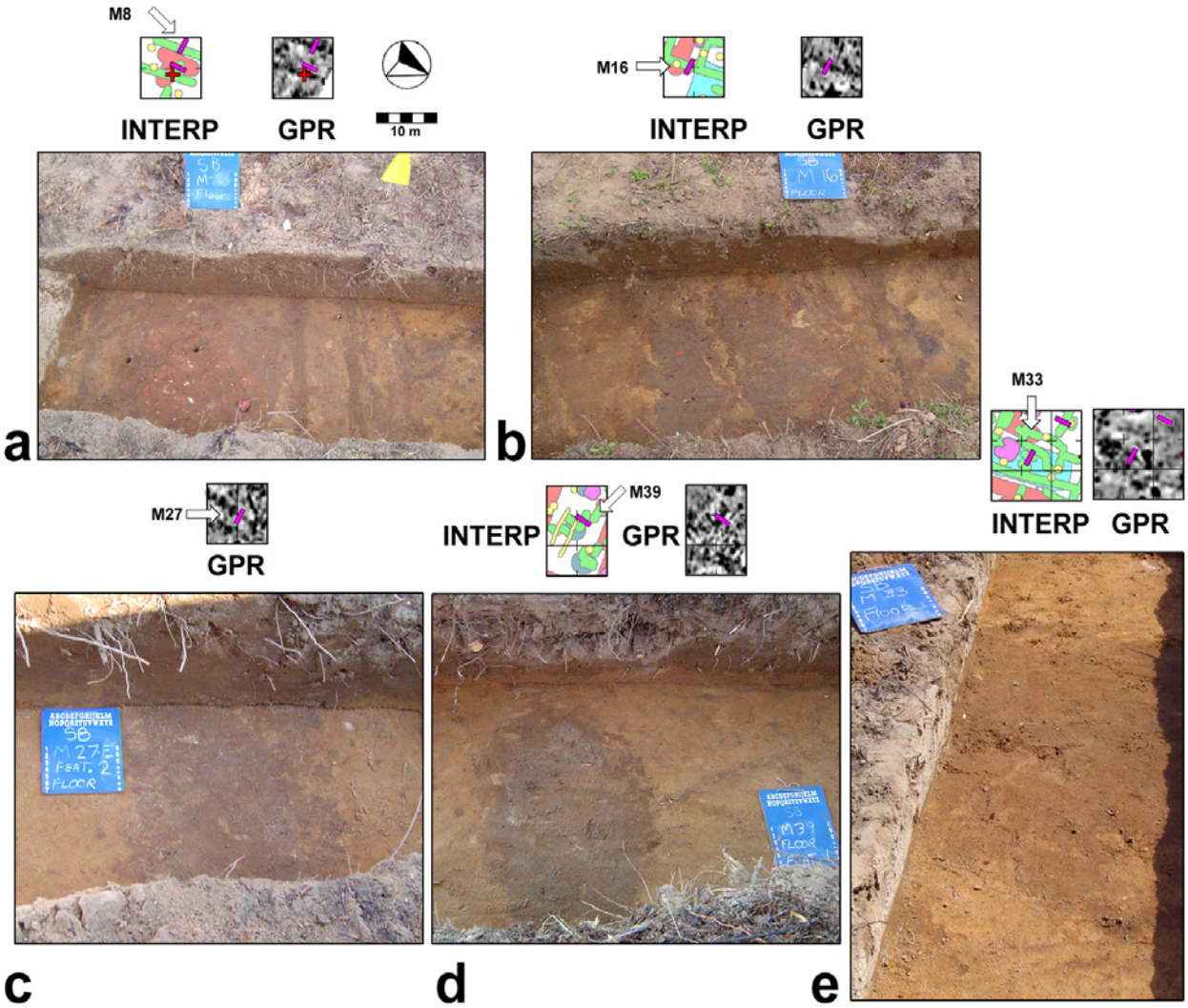


Figure 5.86. *Linear GPR anomalies and their archaeological sources. a,b) Robust linear GPR anomalies and sources in trenches M8 and M16. c-e) Subtle linear GPR anomalies and their sources in trenches M27, M39, and M33. Note plow scarring in (a). KEY: INTERP=interpretations; GPR=ground-penetrating radar.*

One surprise, pointing to the intensity of the site's occupation, was the large number of post holes discovered by the excavations. They occur in two broad types illustrated in (Figure 5.87). One is associated with a rectangular "post pit" of substantially larger size (Figure 5.87a-d). The second includes isolated posts (Figure 5.87e,f). Unfortunately, given their small sizes and relatively low volumes, many could not explain observed and very broad resistivity anomalies, whether positive or negative (those illustrated in Figure 5.87b,d,f all occur at the loci of broad resistivity anomalies). Post holes of all types did explain associated linear GPR anomalies, robust and subtle (Figure 5.87c,e). In fact, close inspection of the GPR depth-slices reveals that many of the indicated lineations are, in actuality, a series of closely spaced point anomalies that probably represent linear arrangements of post holes (see Figure 4.45).

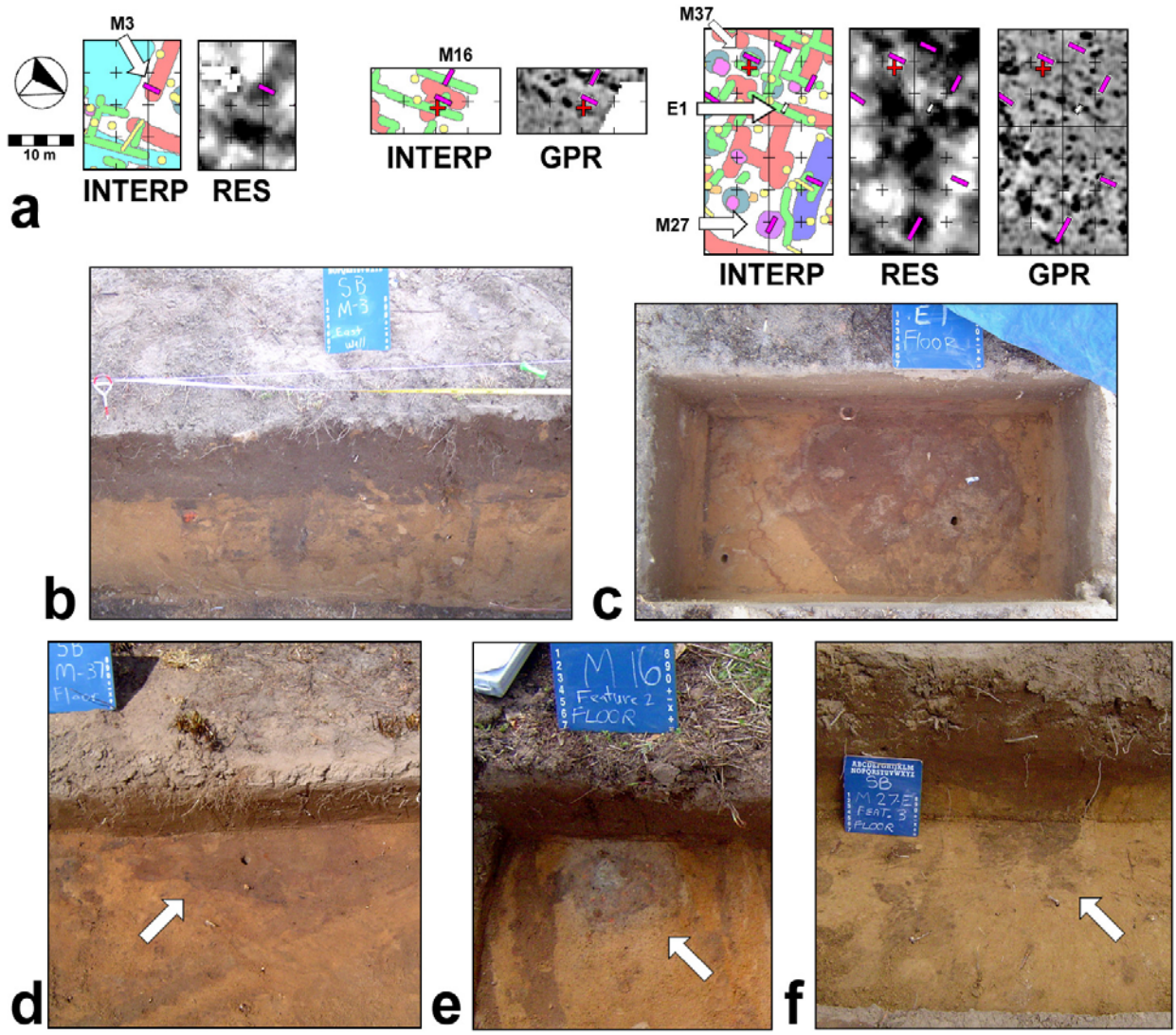


Figure 5.87. *Examples of post holes at Silver Bluff and associated anomalies, when present. a) Geophysical evidence and interpretive maps. b-d) Post holes within associated post pits. e,f) Simple post holes. KEY: INTERP=interpretations; RES=resistivity; GPR=ground-penetrating radar.*

GPR revealed a broad area anomaly that was quite deep, visible in a depth-slice 80-100 cmbs (Figure 4.45e). It was speculated that this anomaly could correspond with a positive resistivity anomaly that was interpreted as a low berm adjacent to an inferred lane, and perhaps indicative of a "deeper" road feature (Section 4.6.3.3). In fact, the evidence given by this low GPR slice in reality depicts a paleochannel (Figure 5.88a) of likely extreme age that extends inland from the nearby Savannah River, as clearly revealed by Trench M42 (Figure 5.88b). Laminar bedding and mineral staining point to the lacustrine nature of these sediments, resulting from ponding.

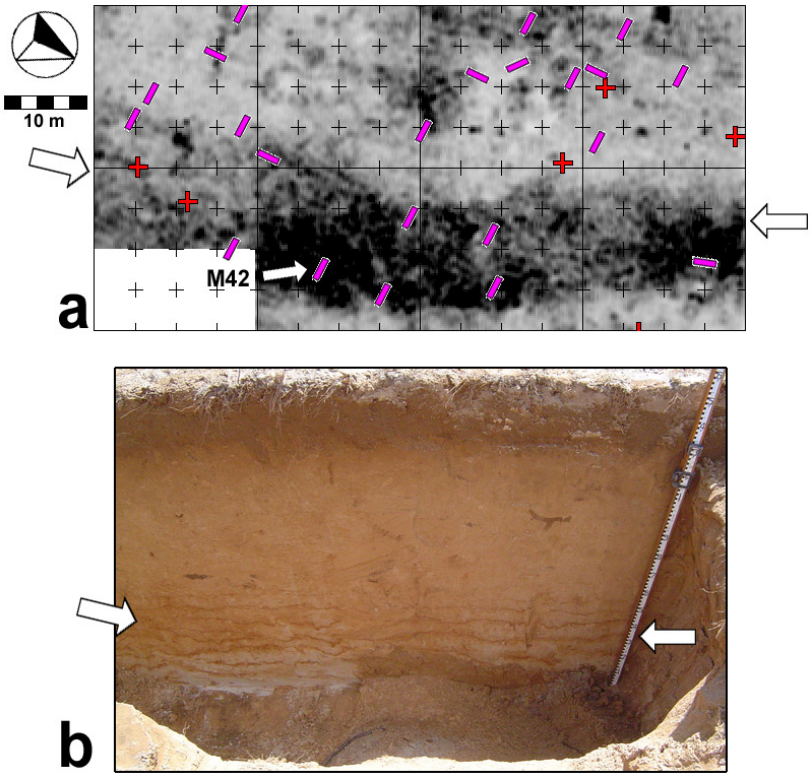


Figure 5.88. *Evidence of a paleochannel in a) a deep (80-100 cmbs) GPR depth-slice (between arrows) and b) the profile of trench M42. Note laminar layers and mineral staining (arrows) that point to lacustrine deposits.*

Accuracy. The accuracy of the remote sensing projections at Silver Bluff is disappointing. For many defined anomalies no plausible source could be located in the excavations. Basic accuracy statistics are given in Table 5.42. They are presented by anomaly types, which are defined by geophysical instrument types. The data are categorized in two ways: (1) whether a clear source was located for an anomaly, and (2) whether the indicated source showed reasonable evidence of being cultural in origin. The latter warrants comment. Many anomaly sources, when located, were indicated only by simple soil stains. While soil stains in an intensely utilized occupation might be argued to uniformly represent culturally placed deposits, so many were amorphous or ephemeral in character that it was difficult for even trained and locally-experienced archaeologists to assign them to meaningful archaeological feature classes. Moreover, such stains could be natural, such as from tree throws or badger dens. A case in point is the paleochannel anomaly of Figure 5.88—a definite source was located in the ground but that evidence was not deemed anthropogenic. The second way anomalies are categorized in Table5.42 is according to whether the archaeological evidence seemed clearly anthropogenic: circular post holes, square post pits, strongly linear features, sharp discrete edges, and compositions filled with artifacts allowed a positive decision in this category.

Forgetting the iron and steel artifacts (which are difficult to miss) and focusing only on the soil anomalies, Table 5.42 shows 34 of the 51 excavation samples yielded a source, or about 67% (95% confidence limits range between 53-78%). Only about half (47%) of the samples excavated could be linked with a definite anthropogenic source, however (95% confidence limits range from 33-61%). By far, the largest source of error

lies in the linear resistivity anomalies, whether positive or negative, where a source could be located for only 4 of 14 (29%).

Table 5.42. *Accuracy assessment of Silver Bluff anomalies by SERDP team.*

Anomaly Type		Anomaly Source Located	Anomaly Source Anthropogenic	Total Samples
1.	*Iron or steel artifact*	20 (100%)	20 (10 historic) (100%)	20
2.	*Subtle magnetic linear*	1 (50%)	1 (50%)	2
3.	*Magnetic point anomalies in linear alignment*	2 (100%)	2 (100%)	2
4.	*Linear negative resistivity*	1 (20%)	1 (20%)	5
5.	*Linear positive resistivity*	3 (33%)	2 (22%)	9
6.	*Area negative resistivity*	3 (75%)	2 (50%)	4
7.	*Area positive resistivity*	1 (100%)	1 (100%)	1
8.	*Subtle linear GPR*	14 (78%)	8 (44%)	18
9.	*Subtle negative linear GPR*	1 (100%)	1 (100%)	1
10.	*Robust linear GPR*	5 (100%)	5 (100%)	5
11.	*Area GPR*	2 (100%)	1 (50%)	2
12.	*Deep GPR*	1 (100%)	0 (0%)	1
13.	*Lightening strike**	0 (0%)	0 (0%)	1
Totals (with #1):		54 (76%)	44 (62%)	71
Totals (without #1):		34 (67%)	24 (47%)	51

*This anomaly type is not cultural in origin and no visual evidence could be located in excavation

Table 5.43 organizes the data by geophysical survey type. Although magnetic gradiometry appears to perform well for soil features, sample sizes are too low for a confident assessment. Electrical resistivity performs poorly for all categories. On the other hand, for GPR a source for an anomaly was located 82% of the time, but the source was determined anthropogenic in only about half the cases.

The apparent difficulty of confidently locating sources for many of Silver Bluff's geophysical anomalies is an issue of concern. Particularly puzzling is the fact that so few of the positive and negative resistivity anomalies could be verified by excavation. It has been suggested that the many lineations seen in the resistivity data are the results of decades of plowing or agricultural practice. This seems unlikely. The width of the anomalies (often 3-5 m) is too great to represent plow furrows and their separation is too

far. This is illustrated by comparing the anomalies in Figure 5.89a,b with the spacing of historic furrows seen in excavation (Figure 5.89c) and in a near-surface GPR depth-slice that clearly illustrates the same furrows (Figure 5.89d).

Table 5.43. *Accuracy assessment of geophysical survey types at Silver Bluff.*

Geophysical Survey Type	Anomaly Source Located	Anomaly Source Anthropogenic	Total Samples
*Magnetic gradiometry**	3 (75%)	3 (75%)	4
Electrical resistivity	8 (42%)	6 (32%)	19
GPR	23 (82%)	15 (54%)	28

*Iron or steel artifacts not included

The resistivity anomalies are discontinuous and appear to form numerous right angles suggesting the presence of constructed features, lanes, berms, compounds around structures, activity areas, or perhaps garden plots. This is well illustrated in Figure 5.89a,b where raw and filtered resistivity data are presented. Regular geometric patterns that are a telltale sign of human works literally leap from the imagery. Perhaps most importantly, *both directions of lineations match exactly the many linear features (walls) that have been revealed by excavation*. The geophysics thus extends mapping of the site's structure gained through excavation. On top of that, *no natural process can account for these patterns*. Consequently, we maintain the inference of an anthropogenic origin to many of the linear resistivity anomalies at Silver Bluff and put this forth as yet another case in the project of "invisible" archaeology. There may indeed be physical evidence of this phenomenon in the ground that is only made visible by quantifying subtle changes in soil resistivity properties.

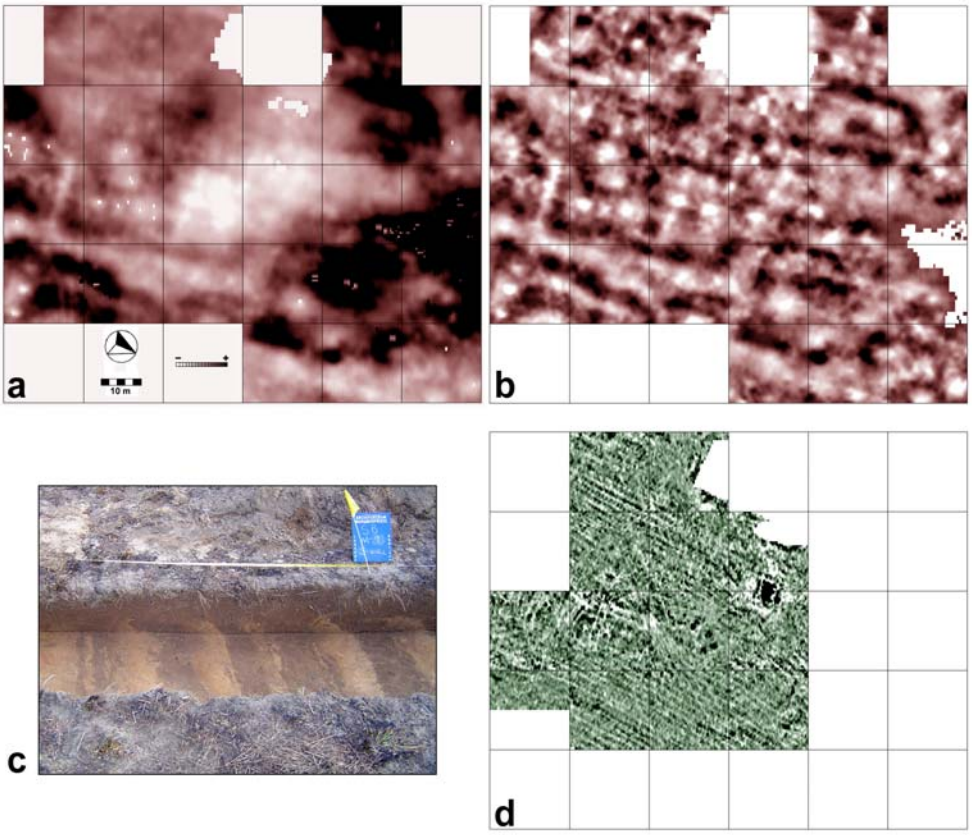


Figure 5.89. *Resistivity anomalies of apparent anthropogenic origin compared to historic plow furrows. a) Raw resistivity. b) High-pass filtered resistivity. c) Plow furrows seen as soil stains in trench M9. d) Plow furrows seen in a near-surface GPR depth-slice.*

5.16.4. Field Investigations at Kasita Town

Michael L. Hargrave, Construction Engineers Research Laboratories (CERL), and Eileen G. Ernenwein, University of Arkansas

Field investigations at Kasita Town were conducted by BHE Environmental, Inc., a CRM consulting firm based in Cincinnati, OH. The BHE crew of five was directed by Dr. Thomas Foster. BHE was selected for the Kasita Town work because Foster was particularly knowledgeable about the historic Lower Creek and had participated in excavations at the site in 2001. BHE's fieldwork occurred from April 11-15, 2005. Kvamme, Ernenwein, and Hargrave arrived several days earlier to mark the locations for the excavation units.

Based on descriptions of past excavations at Kasita Town and similar sites, it was anticipated that the majority of archaeological features would be burials and post molds, and possibly clay extraction pits and other material remains of houses such as daub (see Section 4.6.4). It was also known that these types of features are very difficult to detect because the majority of post molds are of small size, and burials and other similar-sized features do not take an easily identifiable shape in geophysical maps. The anomalies identified for testing included a great many small anomalies that could potentially be burials, clusters of post molds, or perhaps some other archaeological features, but they are also just as likely (or perhaps more likely) to be non-archaeological disturbances and geological perturbations. Specific cultural interpretations were therefore not made.

Instead, it was decided that the field testing effort would focus on exploring the nature of the anomalies so that they might be interpreted afterward (see Section 5.12.7).

Although a total of 41 excavation units (including long 11 backhoe trenches and 30 stripped blocks of 2 x 4 m) were initially planned (Section 5.14.5), some were not excavated. Ultimately, fieldwork consisted primarily of the mechanized excavation of 26 stripped blocks and nine backhoe trenches (Figure 5.90). At the last minute the equipment subcontractor was only able to provide a backhoe with a 60-cm wide bucket, which dramatically slowed the rate of excavation. Many of the stripped blocks were also narrower than planned, with lengths ranging from 5.5 to 14 m. In addition, a few of the units had to be excluded from the final analysis because they were not excavated deeply enough to adequately test the targeted anomaly. Many of the units were quite large, however, so they effectively exposed the sources of numerous anomalies, including ones not identified by the geophysical surveys. No shovel probes, auger tests, or hand excavation units were utilized at Kasita Town. An Oakfield soil probe was used to investigate the depth and fill characteristics of a number of the features. All of these factors ultimately resulted in the excavation of 35 units and testing of 48 of the anomalies predefined in Section 5.14.5. The realized distribution of anomaly types is summarized in Table 5.44.

a **b**

Figure 5.90. *Excavations at Kasita Town. a) BHE crew documenting the stratigraphy in a backhoe trench, April 2005. b) BHE crew troweling a stripped block. The oval stain is a possible burial pit that was not investigated further.*

Table 5.44. *Summary of Kasita Town geophysical anomalies and excavation units.*

Code	Name	Stratum	Population	SB	BT
1	mag lines	M	11	5	
2	mag lines, subtle	M	16	3	2
3	mag polygons	M	111	7	2
		Subtotal:	***138***	***15***	***4***
4	GPR lines	G	22	7	
5	GPR polygons	G	72	5	1
		Subtotal:	***94***	***12***	***1***
6	res lines	R	5		3
7	res polygons, high	R	7	1	
8	res polygons, low	R	12	3	
		Subtotal:	***24***	***4***	***3***
9	mag ditch, robust	D	6		4
10	mag ditch, subtle	D	5		1
11	GPR ditch	D	4		1
12	res ditch, high	D	4		1
13	res ditch, low	D	4		2
		Subtotal:	***23***	***0***	***9***
		Total:	***279***	***31***	***17***
			Total # of Anomalies Tested:	***48***	

Abbreviations: mag = magnetic gradiometry, GPR = ground-penetrating radar, res = electrical resistance, M = mag stratum, G = GPR stratum, R = res stratum, D = "ditch" stratum, SB = 2 x 4 m strip blocks, BT = 1 x 8 – 1 x 20 m backhoe trench.

The excavations soon revealed two important things about the nature of archaeological preservation at Kasita Town. It became obvious that the survey area had been impacted by grading, presumably when the airfield was built prior to World War Two (see Section 3.3.4.4). Many excavation units uncovered an extensive plow zone that appears to have homogenized the upper .2-.3 m of cultural deposits, such that the upper portions of archaeological features were heavily disturbed. In other words, Kasita Town's living surface had been plowed to the extent that features are no longer recognizable. What remains are features below the plow zone that extend rather deeply into the ground—mainly burials and post holes. This explains why these are the only types of features discovered in the ground by the excavations (see below). Another complicating factor with the site stratigraphy and preservation was the discovery that the land had been graded, probably during construction or maintenance of Lawson Airfield. Some areas had been cut down, with the result that the historic plow zone was missing. Other areas had been filled, as evidenced by a layer of spoil atop the pre-construction humus and plow zone. The evidence of this was quite clear, as some portions of the site had as much as 25 cm of overburden on top of the plow zone (which includes the remains of the Kasita Town living surface). In other areas a great deal of material had been removed such that some archaeological features (including burials) were found within .2 m of the modern surface and the plow zone had been scraped away. This cutting and filling has impacted but has not completely destroyed the archaeological remains of Kasita Town. Unfortunately, the patchy distribution of spoil and missing A-horizon diminished the potential for reliably detecting geophysical anomalies associated with features. Figure

5.91 shows a photograph of one backhoe trench where the plowed buried living surface and overlying spoil are clearly visible.

Figure 5.91. *Photograph of a portion of backhoe trench 2 (BT2) showing the plow zone and spoil from military grading of the airfield.*

The success of a geophysical survey depends heavily on the preservation of the targeted archaeological features, so the plowing of the Kasita Town living surface was seriously detrimental to the project's efforts. Although it was expected that the Kasita Town archaeological record would be very limited due to the ephemeral nature of Creek Indian architecture, it was also hoped that geophysical sensors might be able to characterize subtle features not typically detectable by traditional excavations. Any hopes of detecting (or discovering by excavation) house floors, pathways, small mounds of daub from ruined walls, or any other archaeological feature left on the Kasita Town living surface were lost, however, when plowing began at the site.

Military grading also greatly complicated the site's stratigraphy and therefore the geophysical results. Three general stratigraphic situations were encountered during excavations: *normal* areas where the plow zone was intact, areas where the normal stratigraphy was covered by 2–25 cm of *spoil* (Figure 5.91), and *truncated* areas where most or all of the plow zone has been removed. These contexts were plotted on a map and it was discovered that they not only reveal two broad zones of cut and fill, but that the cut and fill zones correspond closely with the unprocessed resistivity data. As is typical for archaeological applications, the resistivity data were high-pass filtered to remove broad trends (as shown in Section 4.6.4), which are typically due to geology, but in this case they result from grading. Yet, the unfiltered version of the data shows the western half of the survey area is characterized by high electrical resistance, while the eastern portion is generally lower in resistance (Figure 5.92). The high resistance in the west is probably a response to the lack of top-soil (plow zone) compared to the eastern half. The eastern half also has the addition of spoil, which is much higher in clay content than the local deposits and therefore may have reduced the electrical resistance even more. The close association of electrical resistance and grading makes it feasible to predict the extent of cutting and filling due to grading based on known stratigraphy from excavations and the electrical resistance patterns. These regions are plotted in Figure 5.92.

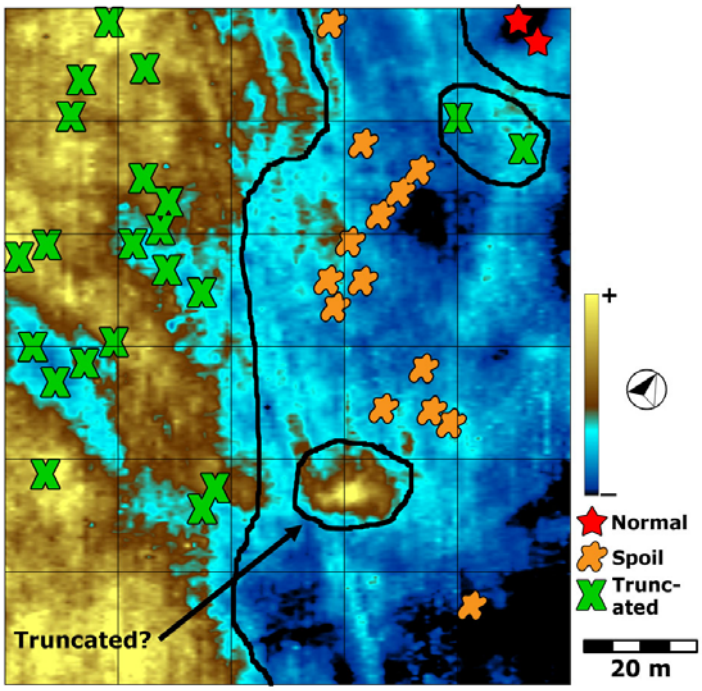


Figure 5.92. *Map of cut and fill areas as a result of grading at Kasita Town, shown by electrical resistance data (background) and excavations (symbols). The western portion of the survey area has been graded causing the Kasita Town living surface (plow zone) to be truncated. Areas to the east have been preserved, but are covered by up to 25 cm of spoil. Only a small area in the northeast corner appears normal, in the sense that it is undisturbed by grading.*

5.16.4.1. Evaluation of predictions by archaeological team
Michael L. Hargrave, CERL

As discussed in Section 5.15.3, the archaeological contractors apparently did not fully understand their responsibility to independently evaluate the degree to which features identified in the excavations could explain the occurrence of remote sensing anomalies. As a consequence, only the SERDP team's assessment of the accuracy of the remote sensing findings is available (see below). Nevertheless, a number of relevant findings may be drawn from the report of their work (Foster 2005). Eighty-three discrete features, many likely archaeological but some possibly natural, were exposed and documented in the mechanized excavations that opened an area of approximately 300 m^2. Clearly, this represents a number larger than the 48 anomalies targeted for the testing program (Table 5.44). None of these features were excavated, so their assignment to functional types is based on their horizontal size, shape, and fill characteristics. Most (68, or 82%) of the features are categorized as post holes. Many presumably provided support for the wall and roof posts of residential structures. Unfortunately, the limited areas exposed by the trenches made it difficult to link individual posts holes and group them as parts of hypothetical structures.

Other feature types included seven likely burial pits, three other pits, two amorphous stains, one daub concentration, one carbon concentration, and a single feature thought to be the plow-disturbed remains of a house wall. This last feature consisted of an amorphous scatter of daub, carbon, and pottery located near several post holes (Foster 2005b:55). Likely burial pits were identified based on their distinctive size and shape, and their presence caused excavation to cease immediately. The variety of archaeological features identified at Kasita is consistent with expectations for a Creek settlement. Feature types associated with structures are expected to include post holes, burials, and

storage pits. Extramural (non-structure related) pits should include isolated post holes, storage pits, and clay extraction pits that were often subsequently used for refuse discard (Foster 2005:55, 2006 in review).

Foster used the results of Gordon Willey's 1936 excavations (which were never fully published), the 2001 excavations associated within a runway expansion, and the 2005 SERDP investigations as a basis for speculating about Kasita Town's overall community plan. He notes that a circular distribution of post holes identified in 2001 could be the remains of a council house. At least two structures are identifiable in Willey's maps, and the SERDP investigations identified three possible structures or structure complexes and a probable midden. Foster suggests that Kasita Town was laid out along the Federal Road, which had long been an important trade route. Feature density in the SERDP study area was much greater than in the area excavated in 2001, and this may indicate that it may be in or nearer to the core area of the site (Foster 2005b:193-195).

In retrospect, Kasita Town was an unfortunate choice for inclusion in this study. Clutter associated with near-surface cutting and filling that presumably occurred when the airfield was built prior to World War Two severely diminished the potential for the geophysical detection of cultural features. Another problem was the unexpectedly low occurrence of relatively large features such as pits. Kasita Town contributed the most to this project as an example of the extent to which surface impacts and feature characteristics influence a site's suitability for geophysical investigation.

5.16.4.2. Evaluation of predictions by the SERDP team
Eileen G. Ernenwein, University of Arkansas

Findings from archaeological excavation at Kasita Town include primarily post holes, burials, and a handful of other archaeological features of unknown type. The findings were classified into four categories: (A) burials, (B) archaeological features, (C) false positives, and (D) no evidence. A total of eight likely burials were found; an example of one is given in Figure 5.93a. Findings categorized generically as features (B) include groupings of post holes (Figure 5.93b), occasional deposits of burned soil, daub, midden, and one possible clay extraction pit. False positives (C) are defined as any target where an anomaly source was located, but was determined to not be of archaeological origin. They include many instances of disturbance by grading and plowing as well as natural sources such as rodent and root disturbances and geological variants. For example, some anomalies were determined to be the result of deep plow scars, such as the one shown in Figure 5.93c.

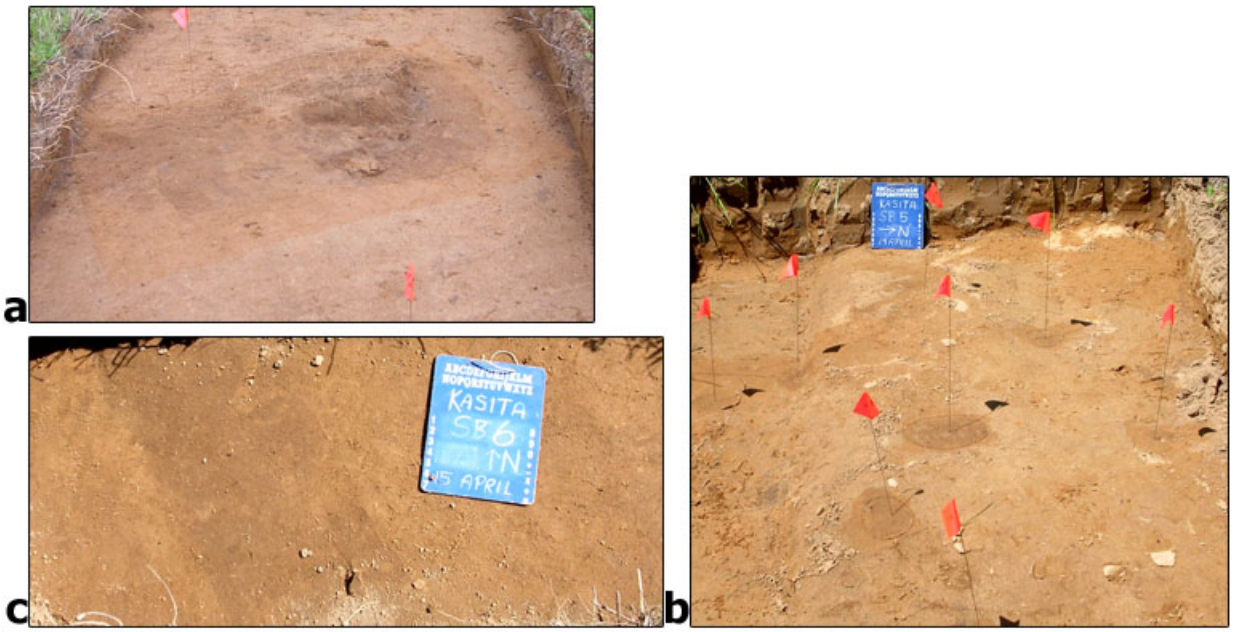


Figure 5.93. *Photographs illustrating the main causes of anomalies at Kasita Town. a) Burial (class A; note that the rectangular shape surrounding the burial is due to moisture collection under a tarp that covered the burial overnight). b) An example of an archaeological feature (class B), in this case a cluster of post molds. c) An example of a false positive (class C), in this case a plow scar.*

One other anomaly type, tentatively labeled as a "ditch" (Section 5.12.7; see also Figure 4.46), was explored extensively with six long backhoe trenches, effectively testing ten different instances of this anomaly type. No observable source, cultural or natural, could be detected in any trench wall or floor despite intensive examination. Since the GPR data indicate this feature to be buried at least 1.6 m below the surface, and the known archaeological features are within the upper meter or less of deposits, it is concluded that this anomaly indicates something unrelated to Kasita Town. The most likely explanation, based on the shape and close proximity to the river, is that these sinuous anomalies represent paleochannels. Since all of the excavations used to test these anomalies had negative results, they are excluded from the remaining discussions.

Table 5.45 summarizes the results of the test excavations and presents percent correct statistics and confidence intervals for each geophysical data type and the overall success rates of the testing. Based on these findings, the chance of discovering a cultural feature when testing a classified anomaly at Kasita Town is about 47%, within 95% confidence intervals ranging between 32-63%. A look at each type of anomaly, categorized by geophysical method, shows that anomalies revealed by magnetic gradiometry were most successfully corroborated by excavation (53%), followed by electrical resistivity (43%) and GPR (42%). The success of electrical resistivity is interesting, because it shows that each positive resistance anomaly (the 3 "line" anomalies and a single polygon) was *not* associated with a cultural source, but all three negative resistance anomalies were found to indicate cultural features. It was also found that area or polygon anomalies were generally more correlated with cultural features than linear anomalies for all data types, probably because the lines were typically smaller.

Table 5.45. *Summary statistics for Kasita Town anomaly testing.*

	n	*cultural?*	*% correct*	*95% Conf Interval*	
01: Mag line	5	2	40		
02: Mag line, subtle	5	2	40		
03: Mag poly	9	6	67		
subtotals:	19	10	**53**	**32**	**73**
04: GPR line	7	3	43		
05: GPR poly	5	2	40		
subtotals:	12	5	**42**	**19**	**68**
06: Res line	3	0	0		
07: Res poly, high	1	0	0		
08: Res Poly, low	3	3	100		
subtotals:	7	3	**43**	**16**	**75**
Totals:	38	18	**47**	**32**	**63**

Abbreviations: mag = magnetic gradiometry, GPR = ground-penetrating radar, res = electrical resistance, n = sample size, culture? = indicates how many anomalies in the category were found to indicate cultural features (related to Kasita Town [classes A and B]). 95% confidence intervals were calculated using an asymmetric approximation. Caution is advised when viewing statistics associated with low sample numbers.

5.16.5. Summary and Discussion

Michael L. Hargrave, Construction Engineers Research Laboratories (CERL)

Excavations designed to ground truth predictions about the occurrence of features based on geophysical data were conducted at each of the four SERDP sites. One goal was to secure an independent, objective assessment of the geophysical interpretations by individuals who were not otherwise involved in the project, and who had substantial expertise in the local archaeology. This was achieved for three of the four sites. At Kasita Town the consultant's field director, an expert in relevant archaeological and historical areas, failed to record the first-hand field observations needed to allow him to assess the potential causal relationships between geophysical anomalies and archaeological deposits. The SERDP team also recorded such data (this was a second goal of the field program), but no independent assessment is available for Kasita Town.

Completion of the field validation exercise provides an opportunity to reflect on the appropriateness of each of the sites for inclusion in this project. Army City and Pueblo Escondido appear to have been excellent choices in all regards. Both sites provided a broad array of features types and were characterized by no serious problems with clutter caused by recent surface modifications, rubbish dispersal, agricultural activity, or biogenic alterations (e.g., extensive rodent reworking of deposits). There had been some concern that the extensive looting reported to have occurred at Pueblo Escondido would be apparent in the geophysical data. In fact, no geophysical anomalies were interpreted as recent looters' holes, and no evidence for looting was encountered in the excavations. There is no reason to doubt earlier reports of serious looting. It remains unclear whether the SERDP team was very fortunate in choosing a well-preserved

portion of the site, or if recent looter holes (which are not visible on the surface) simply do not show up well in the geophysical data.

In contrast, Kasita Town and Silver Bluff each presented challenges to the SERDP researchers, and in retrospect, they were not good choices for test sites. A brief visit to each site would have made the SERDP team aware of the potential difficulties. Previous archaeological investigations at Kasita Town (O'Steen et al. 1997) alluded to localized impacts to the site but noted the coherent patterning in the distribution of diagnostic ceramics as evidence of the site's research potential. Evidence for grading in some areas must have been observed when shovel tests were excavated at 15-m intervals, but this was not reported. To be fair to previous investigators, however, the grading clearly did not severely impact the integrity of the site as a whole, although it did complicate geophysical efforts to detect subsurface features. One might also question the wisdom of selecting Kasita Town because post holes, which represent the most common type of feature, are generally too small to be detected by geophysical methods. However, pit features and burials are known to co-occur with Creek domiciles (Foster 2005b), and those feature types are large enough to be reliably detected. On balance, only the failure to become aware that some portions of Kasita Town had been graded is regrettable.

Two of the limiting factors at Silver Bluff were the amount of previous excavation and the generalized scatter of abandoned pin flags. Here too, a brief site visit would have been useful. In fact, Silver Bluff proved to be the most confusing and frustrating of the four sites for reasons that are not well understood. An examination of Figure 4.45 shows that geophysical anomalies of one type or another literally cover most of the site area. This seemed plausible given that the 18th century trading post's stockade would have constrained many activities, and may have resulted in a more intensive use of the core area than would have occurred if no defensive walls had been present. A truism in archaeo-geophysics is that increasing the area surveyed almost always improves one's ability to interpret the data. At Silver Bluff it would have been beneficial to survey a larger area. For example, this would have indicated whether the linear resistance anomalies that have been speculatively related to plowing are restricted to the immediate area of the trading post. Difficulties in interpreting excavation results from Silver Bluff exemplify another maxim of archaeo-geophysics. Some sites are excellent candidates for geophysical investigation, but others simply are not. Experienced surveyors can make educated predictions, but ultimately a small trial survey is always wise.

At Silver Bluff it was particularly difficult to determine whether features identified in the excavation units were the cause of the targeted geophysical anomalies. Many of the anomalies identified at the site were clearly linear, whereas most of the features were not. Intensive, repeated plowing of soils poorly suited to agriculture may explain many of the linear anomalies, but this explanation must be viewed as speculative (see also discussion in Section 5.16.3). The linear anomalies do not correspond to individual, clearly discernable plow furrows as has been seen at other, recently cultivated sites. Ultimately, the Silver Bluff excavators did not attempt to quantify the verification rates for the various anomaly categories. On balance, while the Kasita Town and Silver Bluff data are not as useful in addressing the SERPD project's basic research questions, they exemplify important issues that confront archaeologists and resource managers who seek to use geophysics in a reliable and cost-effective manner.

It is useful to compare Army City and Pueblo Escondido in terms of the extent to which predictions about the presence of features were verified by archaeological excavations. Validation rates will be of interest to many readers, including CRM practitioners who have little or no previous experience with geophysics and are curious about its potential reliability. It is worth noting that there is a considerable range of variation in the information return and reliability of archae-geophysical surveys done in the U.S. today. This variation stems from several factors, including site characteristics and the expertise, access to the appropriate sensors, and performance of the surveyors. The four SERDP sites exhibit highly variable site characteristics, but the surveys were performed by two teams (the Archeo-Imaging Lab of the University of Arkansas surveyed at Army City and Pueblo Escondido; Archaeophysics LLC surveyed at Kasita Town and Silver Bluff) with substantial expertise and that used very similar equipment and data collection protocols. In short, one can assume that most of the variation in survey results is due to different site characteristics, not variation in the quality of data collection or processing.

The best results appear to have been achieved at Army City. Here the presence of the predicted type of feature was verified in 70% of the excavation units by the archaeological evaluation team, and 76% by the SERDP team. If one also includes situations where a feature other than the type predicted was documented, the verification rate increases to 86% and 84% for the respective evaluations (Section 5.16.1). Here one must keep in mind, however, that the "other feature" was typically described as a soil or fill zone, not necessarily an intentionally constructed facility (Kresja 2005). The linear anomalies associated with streets and street-related features are readily detected in geophysical data, perhaps because the human mind and eye tend to seek out coherent, repeating, and spatially extensive patterns. But the dirt-filled grooves and ruts in a dirt street that are now buried beneath a layer of dirt are understandably difficult to discern within the confines of a relatively small backhoe trench. Even highly experienced field archaeologists rarely, if ever, have the opportunity to excavate such features.

If one sought to characterize Army City's archaeological deposits with a single word, "massive" might be a good choice. The site includes dense concentrations of building debris, including cement (some of it reinforced with iron rods), stucco wall cladding (backed with wire mesh), and iron and fired clay pipes and tiles. Hand excavation at Army City would be extremely expensive, and very large volumes would need to be excavated to discern architectural patterns. Geophysical survey is undeniably the most expedient approach for documenting such large, recent, high contrast sites. This is a significant point in that many roughly similar complexes of barracks and related buildings dating to the first and second world wars are present at U.S. military installations. Some of these are excellent candidates for geophysical mapping; for others, the presence of subsequent components has diminished the potential for geophysical investigation.

Superficially, Pueblo Escondido is comparable to Army City. Both sites were comprised of a large number of buildings whose arrangement reflected a coherent community plan (in the case of former this observation is somewhat inferential). But while Army City's archaeological remains are massive and high-contrast, the remains of Pueblo Escondido are highly ephemeral. TRC's crew of highly experienced archaeologists found it challenging and at times frustrating to identify and document the

limits of architectural remains. This was, of course, not true in all cases. Most of the pit houses encountered in trenches were readily discernable. Surprisingly (given the results of earlier investigations), only two definite adobe features were documented, the base of a wall and a prepared hearth. Possible adobe was observed in several other locations but was very poorly preserved. It is uncertain whether many adobe rooms like those excavated in the 1960s and 1970s (Beckes 1975; Hedrick 1967) are still present in other portions of the site, or if most such surface structures have been destroyed by looting or erosion of the surface. The SERDP work has demonstrated, however, that many pit houses remain intact, and that the site retains tremendous research potential and cultural significance (TRC Env. 2005:76).

At Pueblo Escondido predictions about the presence of specific feature types were verified in 49.2% of the trenches by the archaeological team and 64% by the SERDP team. If one adds the cases where a feature type other than the one predicted was documented, verification rates increase slightly to 50.8% and 70%, respectively (Section 5.16.2). Overall, structural features (floors and walls) were very reliably detected in the Pueblo Escondido geophysical data. This finding has very important implications for archaeologists working in the region. However, each site is a nearly unique combination of the factors that influence the potential usefulness and reliability of geophysics (e.g., contrast, clutter, details of the local environment and archaeological deposits), and work at additional sites is needed to corroborate the success at Pueblo Escondido. On balance, however, geophysics appears to be a promising approach for mapping subsurface architectural features at sites similar to this one.

At Pueblo Escondido the SERDP team was frustrated by the low verification rates for extramural features including lineations and systematic small features. In the geophysical data these anomalies are clearly spatially correlated with the architectural remains. It is conceivable that these anomalies are associated with highly ephemeral remains of features used in activities that occurred outside the houses. A QuickBird satellite image shows that vegetation occurs in several north-south linear bands (Figure 3.4a), and these could be part of a larger-scale pattern in the distribution of the short, east-west oriented room blocks, but this explanation seems unlikely. It is interesting to see that at both Army City and Pueblo Escondido linear features that were quite apparent in the geophysical data were often difficult to detect in the field. It is true that the backhoe trenches used to investigate anomalies were relatively small, but unit size is generally a factor in archaeological excavation. Silver Bluff is an even more extreme example of the difficulty of identifying the cause of linear anomalies. At Silver Bluff the linear anomalies have been speculatively related to plowing (but see Section 5.16.3), but this explanation is certainly not applicable at Pueblo Escondido or Army City. More likely, it is a tendency of human visualization to "connect" discontinuous elements in an image (Avery and Berlin 1992). Anomalies interpreted to be "linear" may, in reality, be a series of discontinuous point anomalies in linear arrangement. Attempting to validate them by narrow trenching is therefore a "hit or miss" proposition. Reinterpretation of Pueblo Escondido by Ernenwein (Section 5.16.2) and some of the physical evidence of post holes at Silver Bluff (Section 5.16.3) suggests this may be the case at both sites.

Several lessons learned in this project will hopefully benefit colleagues who may undertake their own ground truthing efforts in the future. In retrospect, it would have been useful to reduce the number of anomalies investigated in order to allow much more

time to be devoted to problematic anomalies and features. While it was necessary to categorize the anomalies at Silver Bluff and (to a lesser extent) the other sites using descriptive categories, those categories were troublesome for archaeologists who had relatively little understanding of geophysics. It was highly beneficial for the SERDP investigators (all of whom are also experienced archaeologists) to work alongside excavators who were very familiar with the local archaeology. A field validation strategy that did not permit such interaction would result in far less useful information.

5.17. REVISED SITE INTERPRETATIONS

This section addresses what was learned about the project sites through the archaeological evaluations. Its principal focus is to address changes in anomaly classifications based on those findings, to correct errors discovered by the excavations, and to present revised interpretive maps.

5.17.1. Revised Interpretations at Army City

Kenneth L. Kvamme, University of Arkansas

Archaeological excavations at Army City largely corroborated the remote sensing interpretations of Section 5.12.4. At the same time, much new was learned about this site. First, it became very obvious that much of the site is extremely shallow and that large amounts of the town remain despite its dismantlement and physical removal after abandonment. The exact loci of streets were rediscovered, as were several alleys, sidewalks, and the sewer and pipeline infrastructure. It was learned that part of General Street was surfaced in sand. The footprints of buildings were reestablished by locating walls, footings, and floors that gave previously unknown form to many buildings.

Much was also learned methodologically that will aid future interpretative efforts in archaeological remote sensing. Rebar in concrete resembles iron pipes, for example. Ceramic pipes look like iron to sensors. Floors lacking a discrete hard surface are difficult to detect when digging in room centers. Some features are better seen in profile, such as gutter cross-sections, than in plan view.

A good lesson is learning not to over-interpret. Referring to highly magnetic and conductive pipes as “metal pipes” might seem appropriate at first glance, but neglects the possibility that highly magnetic and conductive ceramic pipes may be present. The possibility of surface rubble resembling architectural features should not be neglected—in one case a ceramic pipe lying on the historic surface was interpreted as a wall (Figure 5.50b). In general, small excavations are insufficient for evaluations except for discrete and pronounced targets of metal, pipes, or super-resistant walls and floors (e.g., of brick or concrete). Frequently, success in locating an anomaly’s source is related to how long and hard one looks for it. This, in part, explains why anomaly evaluations by the SERDP team in Section 5.16 usually exhibited somewhat higher validation rates than corresponding evaluations by the archaeologists. The SERDP team simply had more time.

Turning to the original anomaly classification of Table 5.16 and the evaluation statistics of Table 5.28, it was learned that many of the classes were totally or nearly correctly specified. The single anomaly interpreted as a chimney base was defined correctly (Class 1), as were nearly all of the anomalies classified as footings (Class 2),

isolated metal artifacts (Class 3), pipes (Class 4), debris piles (Class 5), and robust walls (Class 6). These anomaly classes are nearly unchanged in the revised interpretive map (Figure 5.94.), except for two. One anomaly formerly classified as a pipe has been reinterpreted as a robust wall (that likely contains metallic components such as rebar), based on reevaluation of the geophysical data. A couple of anomalies originally thought to represent robust walls are now viewed as robust alley gutters owing to their spatial placement. Additionally, one debris pile located in an area interpreted as a sidewalk turned out to contain significant sidewalk slabs in the excavations (Figure 5.65i). Consequently, all "debris pile" anomalies within interpreted sidewalk areas have been defined as a new class named "sidewalk debris" (Class 18; Figure 5.94).

Only 8 of 14 (57%) subtle wall anomalies (Class 7) were determined to be correct as to type in the field evaluations (Table 5.28). To remedy this, the geophysical data were revisited and the subtlest "wall" anomalies were removed from the map (Figure 5.94.). Additionally, the wall anomaly representing a pipe laying on the historic surface was reclassified as a "debris pile."

Anomalies classified as robust floors (Class 8) were correct only 58% of time, but the biggest source of error was associated with anomalies interpreted as subtle floors (Class 9), of which only 20% were correctly specified (Table 5.28). The excavations showed that most floors had been removed in the town's dismantling, with only a substrate such as sand, gravel, or debris remaining, if any evidence could be located. This poor performance called for the removal of most subtle floors in the reinterpretation save for those few with compelling geophysical evidence or validated by excavation (Figure 5.94.).

"Street" anomalies were generally reliable; the single error over the "surfaced" street area (Class 10) was probably due to the very limited exposure (1 x .5 m) of a hand excavation (Table 5.28). All street-associated anomalies, the foregoing, street gutters (Classes 11-12), street ruts (Class 13), alley gutters (Classes 14-15), and sidewalk edges (Class 16) remain nearly unchanged in (Figure 5.94.) except for the few additions to the robust alley gutter class discussed earlier.

Finally, it is emphasized that the supervised classification methodologies pioneered in Section 5.7.3.2 and 5.8.3 offer a promising new methodology for updating interpretive maps based on known archaeological results. In those studies it was shown that several of Army City's feature classes—pipes, concrete, gutters, and "other" archaeological phenomena as a group— could be accurately modeled based on relatively few samples of the classes revealed by excavation. An implication for the future is that it might be possible to model additional locations of specific archaeological feature types as a means to generate interpretive maps (and such maps could even be associated with reliability and performance statistics).

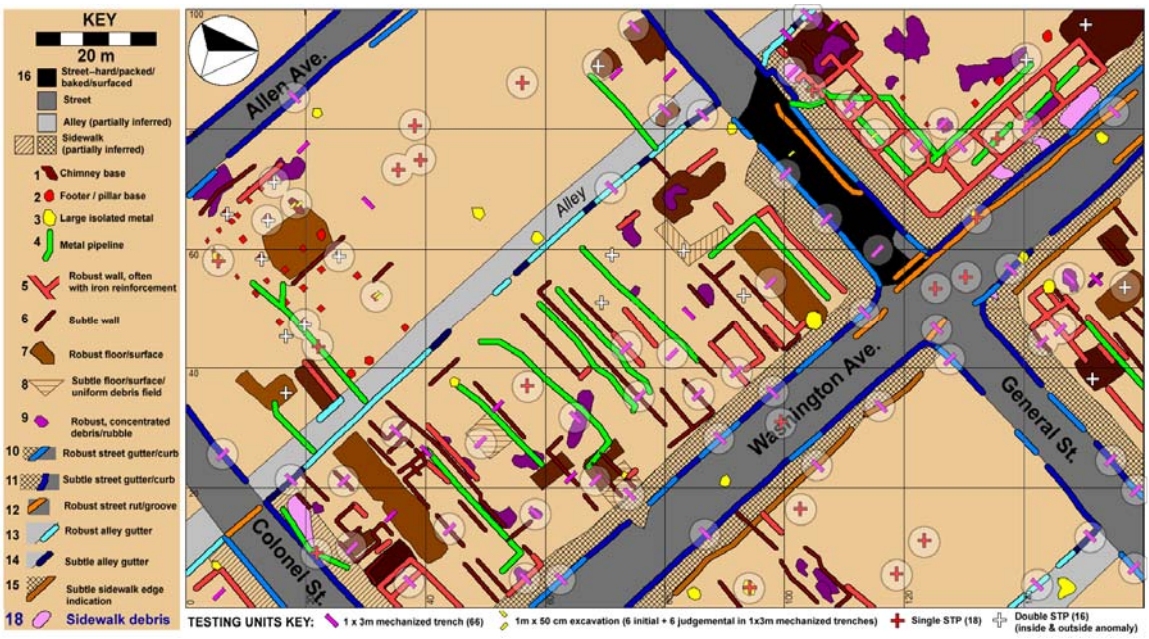


Figure 5.94. *Revised classification of Army City anomalies based on knowledge gained in the archaeological testing program. Circled excavation units indicate correct anomaly classifications. STP excavations not associated with interpreted anomalies represent background tests.*

5.17.2. Revised Interpretations at Pueblo Escondido

Eileen G. Ernenwein, University of Arkansas

Test excavations at Pueblo Escondido helped clarify the nature of the archaeological deposits. The initial interpretation (Figure 5.45) indicated the floors and walls of a series of pueblo room blocks oriented east-west, as well as a few isolated structures thought to be pithouses. Surrounding these habitation structures were dense clusters and lineations of small amorphous anomalies, giving the impression of intensely used space such as patios or food processing areas. These were indicated in the interpretive maps as both small amorphous anomalies and lineations. Many other small anomalies and lineations throughout the survey area were also identified as potential archaeological features. In addition, the magnetic data was used as an indication of which features might be burned or contain magnetic rocks. Test excavations using a random sampling design allowed a statistically unbiased evaluation of each type of anomaly, although sample sizes were generally too low for rigorous conclusions. Prior to the field work it was anticipated that excavations would lead to more specific archaeological feature types, ultimately resulting in a much more complex interpretive map with many feature categories. Instead, excavations successfully verified the presence of many features, but did not help to narrow down the type of anomaly any more than was already known in most cases. In other words, the excavations proved nearly as ambiguous as the remote sensing results. The final interpretation therefore contains fewer feature categories, but a perhaps clearer and more useful map of the site.

The fourteen anomaly categories of Table 5.17 were examined and found to contain only five distinctly different types of features: (1) floors, (2) burned floors, (3)

adobe walls, (4) features, and (5) burned features. Each anomaly was scrutinized and either re-categorized into these five classes or deleted if it could not be verified. This resulted in the removal of all lineations, which were replaced with smaller polygons in some places, a few small anomalies, and two floors. The excavations also led to the addition of several floors, adobe walls, and small features. Moreover, first-time interpretations are also given of the .2 ha area of geophysical data that was acquired while test excavations were taking place.

The new anomaly class called "floors" included the majority of anomalies from old categories one, three, and four (room block floor, inferred room block floor, and outlying floor) as well as a few from old categories six and seven (room block walls, robust and subtle) (Table 5.17). The majority of predicted room block walls turned out to be the edges of pithouses (see Figure 5.73). They did not much differ from floors so they were lumped into the floor category. One new piece of information learned by test excavations at Pueblo Escondido was the fact that the majority of "room blocks" were not typical pueblo constructions. Instead, many were constructed like pithouses. Yet they were arranged much like a pueblo, which is why they were interpreted as room blocks. Now we know that the habitations structures are better described as "pit-rooms" and represent a transitional form of architecture between pithouse and pueblo. The main difference between these two forms of architecture is that pueblos are constructed mostly above ground and neighboring rooms share a wall made of adobe. In contrast, pithouses have floors that are at least partially sunk into the ground, with earth forming the lower part of the walls and the upper parts constructed of a variety of materials. The pit-rooms encountered during excavations would have had a combination of these elements including a partially sunken floor, but some walls constructed of adobe and most rooms organized in neat rows.

Burned floors were a second new classification of anomalies, and included most of the old categories two and five (magnetic room block floor and wall) as well as one of the burned interior room features (category nine) that could not be differentiated from the floor itself (Table 5.17). In general, many of the interior room anomalies previously identified were difficult to differentiate from a structure's floor; many of them were simply incorporated into the new "floor" class. In hindsight, many of these interior anomalies are simply clusters of debris inside a structure, so their boundaries are indistinct. Additionally, many excavations were halted as soon as a feature was uncovered, so the floor was sometimes left buried and the exact nature of the anomaly was not determined. The exception to this is the discovery of a few hearths (see Figure 5.75 and 5.76), which are more easily identified because the burning makes them visible to the eye and their boundaries are discrete.

Two adobe walls were encountered during excavation, but in general most of the predicted walls were the edges of pit-rooms or pithouses. Because adobe walls are distinctly different from the edges of pit structures and they are an indication of architectural style, they were classified into their own category. A few additional adobe walls were identified in the data set based on the knowledge gained from excavated these two.

The last two new anomaly classes, "features" and "burned features," encompass the remainder of the old anomaly categories. Features included the majority of anomalies previously in categories twelve and thirteen, and also some of the lineations from

categories ten and eleven (Table 5.17). These classes include many anomalies verified by excavation, but whose function could not be specified. Although a handful of features could be identified as something more specific, such as a storage pit, most were unidentifiable so they were put into the general "feature" category. Many of the lineations could not be verified by excavation, but those that were confirmed the suspicion that they were composed of closely-spaced discrete features (see Figure 5.75 and 5.76). The new "burned feature" category is very similar to the "feature" category, but these features are all known or suspected to be burned in some way (and possibly incorporating magnetic stone such as fire-cracked rock). This new class includes several hearths, and many other features containing ash, charcoal, burned wood and pottery, and fire-cracked rock.

Figure 5.95 shows the final interpretation of Pueblo Escondido. In general, it appears much like the original interpretation (Figure 5.45), but it is simplified and represents a much more thorough understanding of the site as revealed through limited test excavations.

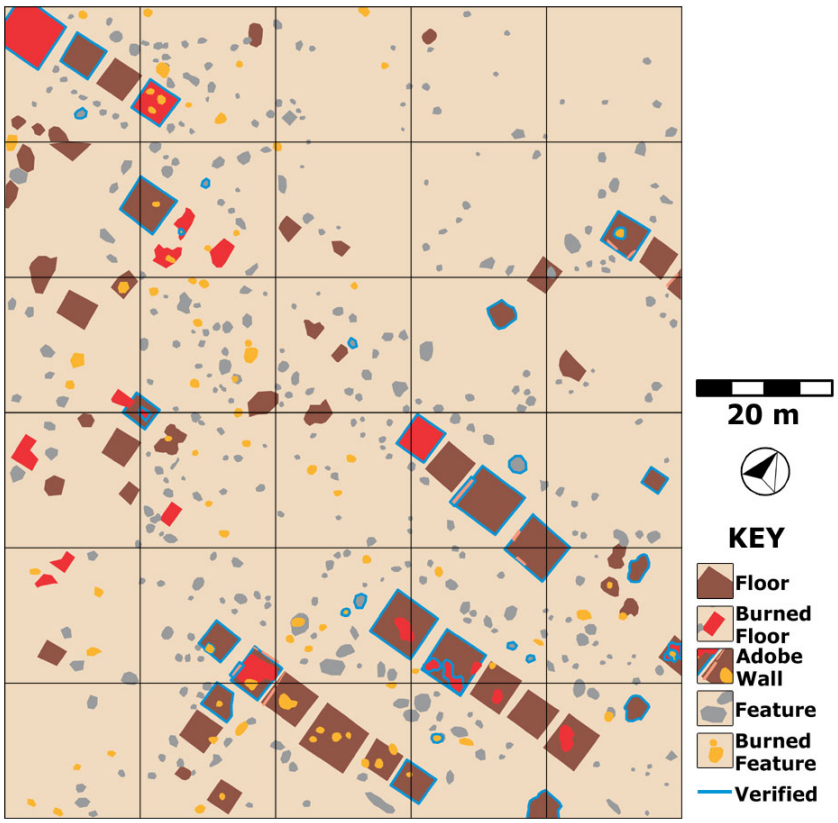


Figure 5.95. *Reinterpretation of Pueblo Escondido based on testing of geophysical anomalies.*

Anomalies that were not verified have been deleted, and those that were confirmed are highlighted in blue. The previously identified pueblo rooms and walls are now known to be mostly pit-rooms and pithouses, and the linear anomalies have been converted to smaller, discrete features. Yet, the fact that this map remains very similar to the initial interpretation shows that the use of remote sensing was highly effective. The combination of multisensor survey and limited test excavations has produced a definitive map of 1.2 ha of Pueblo Escondido. In addition, the findings show the research potential

of the site, with rich cultural deposits. These investigations have only examined a limited portion of the site, which retains potential for future work involving the investigation of the nature and timing of architectural and cultural change during the El Paso phase, the spatial patterning of occupation and reoccupation sequences, the detailed study of well-preserved flora and fauna, and the testing of remote sensing strategies. Of particular interest to the SERDP research team is the use of additional types of remote sensing tools at the site, as well as a refinement and anticipated improvements to the overall strategies employed for "ground truthing" of geophysical anomalies.

5.17.3. Revised Interpretations at Silver Bluff Plantation

Kenneth L. Kvamme, University of Arkansas

Revising interpretations for Silver Bluff Plantation was a challenge because of the many vagaries of the archaeological testing program. As discussed in Section 5.16.3, sources for many anomalies could not be located and even when they could archaeological interpretation was frequently unclear. Several of the interpreted anomaly classes defined in Table 5.18 held up fairly well to validation. A source was located for all iron and steel artifacts (Class 1), but the validation demonstrated that only 50% can be expected to be of historic age (Table 5.42). Magnetic gradiometry also defined two classes of soil anomalies: one appearing as a series of points arranged in a long line (Class 3), and the other as short lineations (Class 2). The former was confidently located while the latter was sourced in one of two excavations. All magnetic anomalies are re-plotted (in yellow) in the revised interpretive map of Figure 5.96.

The archaeological testing also indicated three of the anomaly classes should be omitted as non-cultural or of recent age. They are the deep GPR anomaly type (Class 12), which proved to correspond with a paleochannel, the lightening strike type (Class 13), clearly non-cultural (a source for this anomaly could not be located in the ground), and the negative linear GPR anomaly (Class 9), which proved to correspond with a narrow grooved trail in the surface (possibly a game or excavator's trail).

Archaeological field testing also indicated that 100% of all robust GPR anomalies (Class 10) were accurately specified. They are plotted in red in Figure 5.96. The excavations also suggested that at least one GPR anomaly originally classified as subtle should be regarded as robust, and this change is also given in the figure. As discussed in Section 5.16.3, these anomalies generally point to robust palisade trenches or large and wide soil features. One of the two "area GPR" anomalies (Class 11) that were excavated revealed a large soil stain (perhaps representing a floor, midden, or activity area), so this class is maintained in the revised map with the negative result omitted (Figure 5.96). The subtle GPR anomalies (Class 8), the most numerous defined type, were more problematic. While a source could be located for 78%, only 44% appeared to derive from certain anthropogenic origins (Table 5.42). Most corresponded with linear soil stains or closely-spaced post holes, both indicative of walls. Those lacking physical evidence are removed in the revised map, as are several of the subtlest ones as a cautionary move (Figure 5.96). It is of interest that the magnetic point anomalies in linear alignment (Class 3) appear to coincide spatially with a subtle GPR anomaly, adding to the confidence of the interpretation (Figure 5.96).

Finally, the electrical resistivity anomalies must be confronted. As discussed in Section 5.16.3, it was difficult to locate a source for most of these anomalies, and few could unambiguously be claimed to be cultural in origin. It is observed that the negative or low resistivity anomalies, shown in black in Figure 5.96, generally coincide with and parallel many of the GPR lineations, particularly robust GPR anomalies. This is potentially significant because all of the robust GPR indications were associated with massive archaeological features (very dark soil stains and lineations), meaning that low resistivity generally corresponds with these areas as well. Admittedly, only 3 of 9 excavated negative resistivity anomalies *not* associated with GPR proved to point to an anthropogenic source, but all were very strong archaeological features. Moreover, the positive resistivity anomalies, shown in white in Figure 5.96 (see also Figure 5.89, an inverted image portraying positive resistivity), could also represent significant cultural features, although only 3 of 10 excavated was associated with a promising cultural feature (Table 5.42). Owing to a lack of confidence in defining what is significant in resistivity (positive, negative, or both kinds of anomalies), these data are not vectorized in this reinterpretation (Figure 5.96).

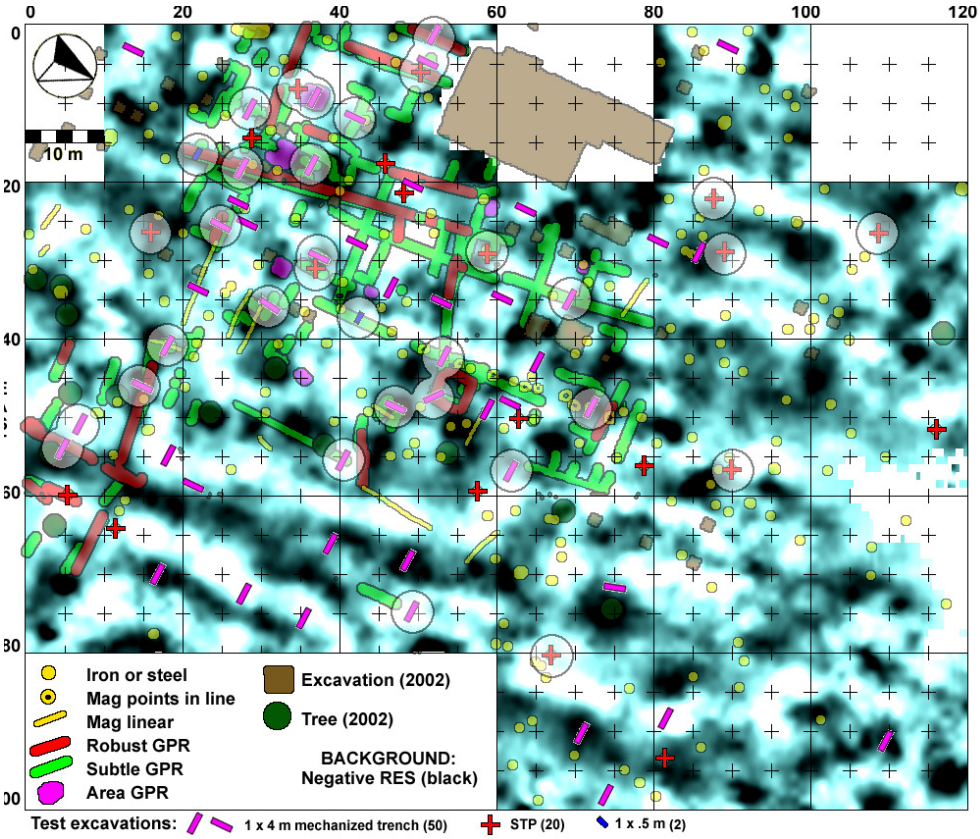


Figure 5.96. *Revised classification of Silver Bluff anomalies based on knowledge gained in the archaeological testing program. Ambiguous resistivity data (background) are not classified. Circled excavation units indicate correctly identified anomalies.*

5.17.4. Revised Interpretations at Kasita Town

Eileen G. Ernenwein, University of Arkansas

Test excavations conducted at Kasita Town were used as a way to explore the nature of the archaeological features that the geophysical anomalies might represent, rather than to test archaeological predictions. Prior to excavation it was anticipated that

the majority of archaeological features at Kasita Town would be post molds, burials, and perhaps clay extraction pits and house material remains such as daub. Excavations confirmed this notion, and also showed that the site has been subjected to major disturbances in the recent past. Not only has plowing destroyed the Kasita Town living surface, but also land leveling seems to have greatly complicated the geophysical data and the ability to detect cultural features. It is likely that many archaeological features were obscured by additions and subtractions of overlying soil and the addition of foreign fill material. Another major discovery was that the "ditch" anomalies are almost certainly not associated with Kasita Town. Their origin was never determined in the field, but they are probably paleochannels associated with the Chattahoochee River. The re-interpretation of the geophysical data based on knowledge of test excavations included the organization of results into two classes of archaeological features: "burials" and "features," the latter of which included mostly archaeological features whose function could not be ascertained by the limited excavations.

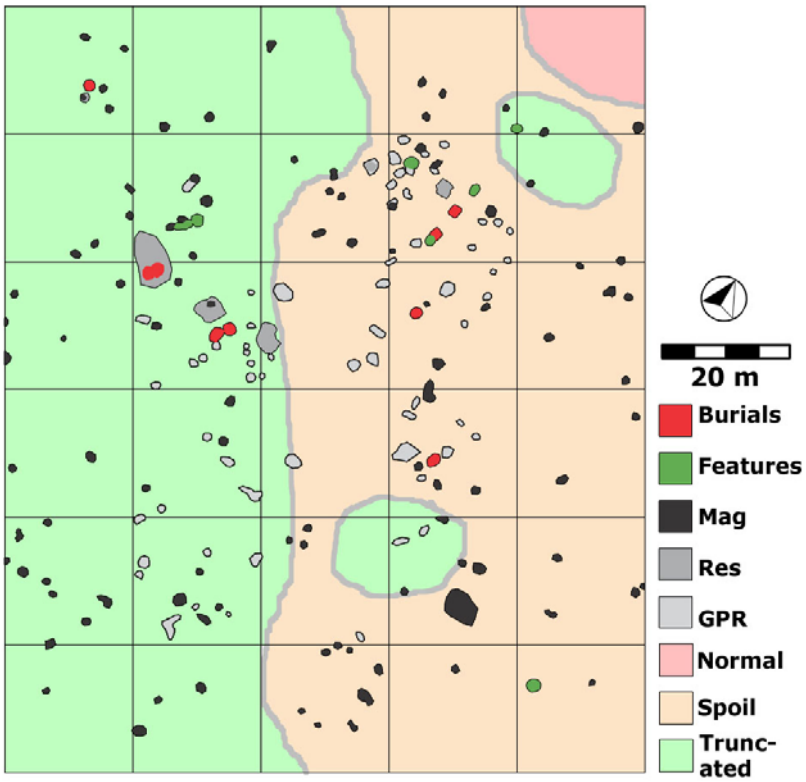


Figure 5.97. *Reinterpretation of Kasita Town based on testing of geophysical anomalies. The two main types of cultural features discovered during excavations, burials and features are shown in red and green, respectively. Magnetic gradiometry (Mag), electrical resistance (Res), and ground-penetrating radar (GPR) anomalies from the initial vector classification are also shown, but those that were tested and did not reveal cultural features have been removed.*

Figure 5.97 shows a final map of Kasita Town geophysical anomalies that have been changed to reflect knowledge gained from test excavations. The confirmed burials and features are shown in color, and the remaining untested geophysical anomalies are presented in gray. In addition, the effects of grading are shown based on the map in Figure 5.92. A principal difference between the new map and the initial interpretation (Figure 5.47) is the removal of the "ditch" anomalies, as well as numerous others that were found to be associated with non-archaeological deposits. It is also clear that all but two of the features and burials discovered are located in the northern portion of the

survey area. One interpretation is that the northern area represents a richer portion of the site, or an area with better preservation. It must also be considered, however, that the northern portion had more geophysical anomalies that were identified for testing and therefore more excavations (more than twice as many). With the high level of disturbance that has occurred it is feasible that some portions of the site have been completely removed, while others could be deeply buried.

6. CONCLUSIONS

Kenneth L. Kvamme, Michael Hargrave, Eileen Ernenwein, Deborah Harmon and W. Fredrick Limp

6.1. BENEFITS OF REMOTE SENSING IN ARCHAEOLOGY

Remote sensing is relatively new to archaeology, but has begun to make impacts on the goals and practice of the discipline. As exploding populations result in massive changes to the landscape, efficient and cost-effective methods to locate, map, and acquire information from sites of our cultural heritage are needed before they are lost. Costs of archaeological excavations have skyrocketed and they commonly examine only trivial areas. Remote sensing techniques allow large regions of the subsurface to be rapidly investigated at relatively low cost. These methods can detect archaeological elements unseen on the surface, precisely map them, and offer interpretations based on form, distribution, context, and measurement characteristics. It is increasingly being realized that remote sensing may offer the only pragmatic means for locating, mapping, and inventorying much of the world's archaeological resources. Although this implies that remote sensing is only a descriptive tool, it also offers many methodological avenues for the *interpretation* of that record, a fact emphasized in this project.

The traditional view of remote sensing sees it as a means to (1) discover new archaeological sites and site features, (2) guide expensive excavations to archaeological features of interest, and (3) produce cost savings by making site explorations more efficient. All this is accomplished non-destructively because remote sensing is non-invasive, permitting the resource to remain intact (an important consideration when exploring culturally sensitive burial, sacred, or ceremonial sites) (Weymouth 1986; Wynn 1986). Remote detection of cultural features can reduce amounts of excavation, and therefore costs, because field teams can go directly to features indicated. Smaller artifact collections may also be recovered, decreasing curation charges because the volume of excavation is reduced when archaeological features are reliably located. Moreover, prior detection of subsurface features lessens the risk of inadvertent discoveries of cultural resources and attendant mitigation costs (Hargrave 2006).

Success in these areas has promoted the notion that remote sensing is suited only for discovery purposes, but recent advances have allowed these technologies to evolve into new domains of inquiry. In some cases remote sensing can yield *primary data* suitable for the study of cultural forms within archaeological sites and landscapes. In other words, culturally patterned anomalies can reveal organization and structure within settlements or larger spaces. The identification and examination of relationships between such individual site components as houses or house clusters, lanes, dumping grounds, public structures, storage facilities, gardens, plazas, fortifications, cemeteries, and the like, can be made through interpreted imagery. Likewise, inter-settlement comparisons of form, estimates of numbers of houses, average house sizes, or the examination of house shapes, orientations, and arrangements of interior components can also be attempted through remote sensing alone (Bales and Kvamme 2005; Gaffney et al. 2000; Kvamme 2003). This capability offers the possibility of direct study of settlement content and form through remote sensing. The compilation of libraries of site plans of Neolithic enclosures and Roman villas in Britain, derived entirely from geophysical data sets, exemplifies this

capability (*The English Heritage Geophysical Survey Database*, http://www.eng-h.gov.uk/SDB).

The ability to image large areas of the subsurface has profound implications when contrasted with traditional methods of archaeological excavation. Their costs and time requirements mean small exposures limited to a few square meters, which forces a mindset that commonly views human spaces in terms of tens of *meters* as opposed to the tens of *hectares* within which most people live and interact. In other words, human activities occur within settlements and cultural landscapes covering hectares or even square kilometers. Until remote sensing, there has not been a ready means to visualize subsurface cultural distributions over such large areas, and there is generally a tremendous amount of archaeological ignorance about the size, structure, and layout of settlements, inter-settlement relationships, and cultural uses of landscapes. Areas of the subsurface "exposed" by remote sensing are several orders of magnitude larger than can be achieved by excavation.

Many of these ideas have been promoted in this project. The remote sensing programs have revealed the "big picture" about several archaeological sites by defining much about their layouts and overall structures. The success of this study rests in application of several innovative approaches. (1) Multiple complementary remote sensing techniques have been applied to a series of archaeological sites, each offering somewhat different visualizations of the subsurface. (2) These results have been integrated to allow simultaneous visualization of multiple dimensions of the subsurface, allowing improved understanding of archaeological content. (3) Large areas have been investigated at each site to facilitate recognition of cultural elements through pattern recognition principles. (4) Interpretations based on remote sensing have been archaeological tested for accuracy, allowing revisions of interpretations.

6.2. BENEFITS OF DATA FUSION

The integration or fusion of multiple remote sensing data sets offers a number of benefits to interpretation and the understanding of archaeological sites. They impart a more holistic view of the subsurface by showing all results simultaneously. A single survey might reveal only part of a buried building, but integrated information from several surveys could illustrate the entire structure as well as its interior components. This aids understanding of single features, but also overall site structure, layout, and organization. Surveys of large areas aid interpretation by allowing imaging of complete features and their forms—a house is better recognized when viewing its full shape as opposed to only a small portion of it. Large-area surveys show context and associations between features further aiding interpretation. By visualizing data from several sensors simultaneously, an improved understanding of sensor relationships, redundancies, and underlying dimensionality is achieved. This is accomplished visually, and with some methods quantitatively through correlation coefficients, loadings, factors, and other statistical indicators.

Graphical solutions for data integration are easy to implement and effectively combine information from disparate sources into interpretable displays. They allow complex visualizations of the subsurface, but their weakness rests in relatively low dimensionality—only 2-3 data sources may effectively be represented. Moreover, these methods are purely descriptive, yielding only images, not new data that may subsequently

be analyzed. Discrete integrating methods, on the other hand, allow application of readily available Boolean operations to any number of geophysical data sets. A shortcoming is the binary maps upon which these methods are based that rely on arbitrary thresholds to define significant anomalies, while more subtle ones must be ignored. Continuous data integrations can yield insights beyond the capabilities of other methods. Robust *and* subtle anomalies may be simultaneously expressed, producing composite imagery with high information content. Interpretive data are also generated in the form of principal component scores, factor loadings, regression weights, or knowledge-based rules that add to understanding of interrelationships and underlying dimensionality. Supervised and unsupervised classification methods are noteworthy because they introduce a predictive aspect to the integrating process. Resulting data fusions, based on multivariate relationships between a suite of sensors, may actually extend the capabilities of subsurface remote sensing because mappings will contain more information about particular anomaly types (or corresponding archaeological feature types) than any single sensor. They therefore offer a possible means to augment prospecting capabilities. Which fusion methods work best may depend on purpose. Some yield visually pleasing results that well-integrate available information while others appear less revealing but offer interpretive or predictive potential.

Although integrated data yield several benefits, when generating interpreted maps the analyst must nevertheless return to the primary data to define boundaries of individual anomalies and determine by which method anomalies were identified. These are important issues in forming accurate interpretive maps (as in Section 5.12). With data fusion some anomalies may grow larger in extent making it difficult to pinpoint the locus of an anomaly's source. Moreover, to interpret an anomaly one must often look beyond its form. Knowledge of the type of sensor that revealed it is critical. If an anomaly is indicated by magnetometry an interpretation of a burned feature may be warranted; the same form of an anomaly may be interpreted as a cluster of stones to electrical resistivity. Without this primary information only limited interpretation is possible.

6.3. TOWARDS AN EPISTEMOLOGY OF ARCHAEO-GEOPHYSICAL FIELD VALIDATION

While geophysics is now more widely used by archaeologists in the U.S. than ever before, it is not yet well integrated into the practice of university based research or cultural resource management (Hargrave et al. 2002). Field validation of anomalies by excavation, or "ground truthing," is even less well established. Important issues such as appropriate sampling strategies and the reliability of alternative field methods relative to site characteristics are only beginning to be considered (Hargrave 2006). It is therefore useful to conclude by discussing some of the issues that will need to be considered as archaeologists begin to formulate an epistemology of geophysical validation.

The factors, that were considered in evaluating the suitability of sites for geophysical investigation and inclusion in this study—contrast, clutter, and environment (Section 3.3.1), are also important considerations when attempting to validate anomalies through excavation. Characteristics of the cultural deposits must be added as another (perhaps the most) important factor. An important point exemplified by the SERDP project (particularly by the Silver Bluff site) is that the acquisition of seemingly "good"

geophysical data does not guarantee that excavation will yield satisfactory validations of interpretations.

Familiarity with the local archaeology is an important component in effective ground truthing under any circumstances, and particularly in situations where features exhibit little visual and textural contrast with their surroundings. Where archaeological excavation is the primary means of validation, the visual and textural contrasts between a feature and its immediate surroundings assume primary importance. Materials commonly associated with historic features such as brick, concrete, and metal almost always exhibit a marked contrast with the surrounding soil in terms of texture, hardness and, in many cases, color. Visual and textural contrasts between an earth-filled pit (the most common type of prehistoric feature in the United States) and the surrounding soil can range from weak to strong. Two factors that contribute to this variation in contrast are the relative abundance of organic debris at the site when the feature was used and eventually filled, and the extent to which site environmental conditions resulted in the preservation of the organic content of the feature's fill. Efforts to identify discrete features in sandy soils where organic stains are faint or absent (as was sometimes the case at Pueblo Escondido), or in situations where midden-filled pits occur in a matrix of midden both demand extensive familiarity with the range of variation in feature types that occur in the region.

A close cultural link with the archaeological deposits being investigated also increase the effectiveness of ground truthing. Army City, for example, is similar in many ways to existing, older sections of towns in the United States. While most archaeologists do not recall the time when dirt-streets were common, the spatial relationships among streets, alleys, sidewalks, and commercial buildings seen in historic photographs of Army City are very familiar. The photographs themselves are invaluable aids to the interpretation of geophysical anomalies (e.g., Figure 3.2). In some cases, the actual sidewalk edges, curbs, and gutters visible in photos of Army City may be those detected in the geophysical imagery. The SERDP investigators' cultural link to the 18th century Silver Bluff site was far more tenuous. Nevertheless, the overall layout of Galphin's trading post, as known from historical descriptions, previous archaeological investigations, and the geophysical data, was far more familiar to the excavators than was the case at Kasita Town or Pueblo Escondido.

Sites that exhibit strong patterning in feature distribution are also more amenable to effective field validation. Army City, an early 20th century "planned" community, epitomizes this observation. At Army City sidewalks consistently occurred between streets and buildings, utility pipes were located in alleys and beneath structures, building entrances were a standard distance from the street, and so forth. These planned patterns are documented in the town plat and historic records (Rion 1960), can be verified through examination of the historic photographs (Hargrave et al. 2002; Kresja 2005; Rion 1960), and are apparent in the geophysical data. Minimal excavation is required to further verify these patterns at Army City. Substantially more excavation would, of course, be needed in situations where the focus was on evaluating site condition and integrity, particularly if subsequent occupations had impacted the archaeological remains of Army City.

Kasita Town exemplifies the difficulty of interpreting anomalies in a site where little patterning is apparent in the geophysical data (here we ignore a few prominent linear anomalies associated with recent drainage pipes, etc.). Anomalies at Kasita Town were amorphous, highly variable in size, shape, and amplitude, and not easily assignable

to standard archaeological feature types. It is now known that much of this variation is the result of localized cutting and filling of the site as it was leveled during airport construction. The absence of clear patterning in the geophysical data made it difficult to use excavation results of a small but representative sample of anomalies as a basis for reliably inferring the nature of the remaining anomalies. Some anomalies in the Kasita Town data are almost certainly associated with real archaeological features rather than simple site clutter, but it would probably require the investigation of an impracticably large sample to develop an ability to differentiate features from clutter.

Archaeologists who are not familiar with geophysical techniques often view the results of careful excavation as the "best" evidence for the presence, absence and nature of archaeological deposits. At Pueblo Escondido, for example, many geophysical anomalies categorized as systematic small features, lineations, and in a few cases, structure floors and walls were not verified by excavation. In most cases, the arrangement of these anomalies in obvious linear patterns made their existence virtually certain (Section 5.16.2). Explanations for the failure to detect features or other deposits associated with these anomalies that were suggested by TRC and/or the SERDP team included the possible misplacement of excavation units, the effects of past looting, and the possibility that vegetation is spatially correlated with architectural remains. While each of these possibilities warrants consideration, they should not preclude consideration of a more controversial possibility. It is likely that, at some sites, geophysics can detect features that are largely invisible to the archaeologist's trained eye and hand. Geophysical techniques "work" because subtle contrasts exist in magnetic, electrical conductivity, dielectric, thermal, and other characteristics. During excavation, archaeologists rely almost exclusively on visual (largely color) and textural contrasts. These contrasts often—but not always—co-occur with the geophysical contrasts. In short, skilled excavators can sometimes detect variations in the archaeological record that are not detected by geophysical sensors, but in the same manner, geophysics can sometimes detect phenomena that archaeologists cannot. The existence of seemingly "invisible" features is a troubling prospect, but the possibility should not be categorically denied. Ground truthing strategies simply need to include alternative means of verifying the presence of features that have no visual or textural manifestation. Examples include the use of geophysical techniques that are based on distinct types of contrast, chemical tests, geological particle analysis, etc. (Mandel and Bettis III: 2001; Sherwood 2001).

The existence of deposits invisible to conventional excavation techniques is not an insurmountable problem for cultural resource managers. It is likely that such deposits will be found to occur in a relatively restricted range of site environments, and to be associated with a relatively narrow range of past activities and feature types. Such deposits are unlikely to represent a "wild card" in assessments of a site's eligibility for nomination to the National Register of Historic Places under Criterion D because they will probably be found to be relevant to a very narrow range of research topics (e.g., the spatial patterning of facilities and activities). Archaeo-geophysicists and archaeologists need to be aware of the possibility of such deposits, and to work together to devise innovative, multi-faceted approaches for their verification or negation.

6.4. FIELD VALIDATION OR "GROUND-TRUTHING"

One of the biggest revelations of the project was the general shortcoming of field archaeology for confirming interpretations gained through remote sensing. This deficiency was manifested in several domains. First, the expense of field archaeology is so great that meaningful statistical sampling is not possible owing to relatively low numbers of excavations. Moreover, because archaeology is invasive, large numbers of excavated holes are frowned upon and may not be permitted because they lead to destruction of a site.

Second, archaeology frequently lacks a certain depth of knowledge when it comes to interpreting findings in the ground. In this project, issues were constantly raised concerning what particular archaeological features should "look like." How, for example, should an alley gutter, a subtle floor, or a street rut appear in the soil? Much ambiguity and uncertainty exists in feature identifications within field archaeology such that absolutely certain identifications may not be possible. This was true at each project site, but particularly at Silver Bluff where numerous soil stains were located by excavation as sources for anomalies although identifying what those stains represented (in terms of a particular type of cultural construction or activity) was often not possible. This uncertainty is exacerbated because the nature of archaeological remains changes not only from region to region but also from site to site as culture types, soils, and specific climate conditions change.

Third, the human senses of sight and touch, upon which archaeologists must rely, respond to only a narrow range of physical phenomena. This range may not be wide enough to detect the sources of a significant percentage of anomalies by normal excavation practice. Several examples of this phenomenon exist in the project from anomalies representing houses at Pueblo Escondido to linear and rectangular anomalies at Silver Bluff that could not be validated by excavation. This deficiency suggest the need for a new ground validating tactic that employs field and laboratory methods from soils science and geoarchaeology as adjunct tools for anomaly confirmation.

Fourth, it is unfortunate that in United States higher education the method and theory of geophysics and remote sensing is not commonly taught to students of archaeology. As a consequence, few archaeologists know about this domain or trust it. The immediate impact is that field archaeologists, although understanding the local archaeology, may not know what to look for when attempting to locate an anomaly source in a "ground-truthing" exercise. Had time allowed it would have been ideal to have the field archaeologists at each site work jointly and share their expertise with the SERDP remote sensing team to develop interpretations for each anomaly in the laboratory prior to fieldwork and during the excavations. Time and funding was far too short during the rapidly conducted fieldwork phase of the project, however.

Last, an ultimate goal of remote sensing may be to give very specific interpretations of individual anomalies, such as "adobe wall," "brick floor," "hearth," "storage pit," or "trail." While an ideal, such interpretations may lead to spurious perceptions of error as when pipes classified as metal turned out to be a highly magnetic fired ceramic and a source of error in efforts at Army City. More generic classes that still give meaningful interpretation of archaeological classes appear to be a superior course of action (e.g., using the generic term "pipes").

6.5. GROUND-TRUTHING VS. VALIDATION

Some archaeo-geophysicists in the U.S. are uncomfortable with the term "ground truthing," particularly when it is used in reference to efforts to verify interpretations of geophysical data by means of excavation. Barring instrument or user errors, anomalies indicated by geophysics are factual: *a real physical cause of an anomaly must exist in the ground*. The issue in validation, then, is not whether a source exists for an anomaly, but whether the interpretation of the kind of archaeological or natural feature an anomaly represents is correct. In most cases, a physical source for an anomaly can be determined by excavation, whether it stems from a change in soil type or a construction made of stone. In such cases the term "ground-truthing" certainly applies and may be employed to confirm remote sensing interpretations. Yet, in many instances the methodology of excavation falls short as a mechanism for validation. The most obvious is when the source of an anomaly is not visible to the excavator's senses of sight and touch. In these cases laboratory analyses or sophisticated field instrumentation must be brought to bear to determine an anomaly's source. Another more common shortcoming of excavation occurs when the source of an anomaly is located in the ground, but the nature of the evidence is ambiguous, precluding *archaeological* interpretation of what the discovered feature represents. This circumstance was highlighted in this project by the many soil stains of ambiguous source and function at Silver Bluff Plantation; archaeological explanations for many of them could not be established. This raises the notion of two sources of error in the anomaly validation process: errors in interpretation by the remote sensing specialist and interpretive errors by the field archaeologist. In such contexts it is difficult and probably inappropriate to use one to validate the other. Given these possibilities, traditional excavation methods cannot always impart "truth" to the process of anomaly validation.

6.6. DATA REDUNDANCIES AND COST-BENEFITS

Weymouth (1986) has observed that magnetometry, resistivity, and GPR generally respond to independent dimensions of the subsurface: magnetism, conductivity, and dielectric contrasts. To this might be added thermal properties. This suggests that, at least theoretically, redundancies may not be present in geophysical data. Analyses in Section 5.7 supports this notion where a Pearsonian correlation matrix between six different sensors showed a maximum absolute correlation of only $|r|=.3$ (Table 5.3) and a first principal component captured only 30% of the total variance (Table 5.4). In this sense, most of the geophysical data explored in this project can be considered to yield very different representations of the subsurface and the archaeology it contains. Viewing any of the geophysical mappings in Section 4.6 suggests this to be the case where new anomalies and different results are portrayed in each one.

Nevertheless, in practice a single archaeological feature might exhibit contrasts simultaneously along several physical dimensions—a stone wall is highly resistant (and low in conductivity); it might be composed of igneous rock and thus be magnetic; and it is likely to possess dielectric differences compared to the surrounding soil and thus generate GPR reflections, for example. This perspective can be witnessed in many of the geophysical results of Section 4.6 as well, where certain subsurface features generate anomalies in several different data sets. This was illustrated most forcefully in Figure

5.9c where certain subsurface entities generated anomalies in as many as six different geophysical data sets.

Aerial (or space) methods will always be cost-effective to acquire because of rapid acquisition of scenes over very large areas. This characteristic must be balanced against data content. In this project, few of the air or space data sets yielded data of use or interest.

In general, as seen in Sections 5.12 and 5.16, GPR is probably the single most productive geophysical survey technique at the types of sites investigated by this project. More anomalies, and significant anomalies, were consistently defined by this method than the others, and most of the vectorized interpretations were derived from GPR. GPR data are associated with greater "costs," however. GPR typically requires somewhat more time to acquire data over large areas (although rapid survey technology is constantly improving), and data processing costs in time are enormous. As a rough estimate, approximately 10-50 times more effort is required to process GPR than other geophysical data sets, but this situation is rapidly improving as well.

The utility of other geophysical methods depends much on site type and conditions (see below). Resistivity proved enormously beneficial at Army City owing to concrete and masonry foundations and floors. In the northern Great Plains, on the other hand, magnetometry is the method of choice because most archaeological features are soil features with variable magnetic properties (Kvamme 2003).

EM induction methods probably represent the principal source of redundancy in the project. The quadrature phase yields conductivity data, which theoretically replicates the inverse of resistivity, and the in-phase component yields magnetic susceptibility (but only to a depth of less than 50 cm), a component recorded by magnetic gradiometry. Kvamme (2006) has shown that the conductivity component consistently reveals less detail than corresponding resistivity surveys. At Army City the conductivity survey failed to reveal any of the many concrete floors detected by resistivity, but the towns numerous sewer pipes were well indicated (they were also shown by magnetometry). In-phase EM was also informative at Army City where likely burned areas and street gutters were revealed and at Pueblo Escondido where numerous lineations and architectural features were defined.

6.7. EFFECTS OF ENVIRONMENT ON DETECTION

Determining the effects of environment on the ability of remote sensing to detect subsurface archaeological features is a difficult issue to confront with only four principal sites targeted in this study. Based on the SERDP teams' experience in more than 25 states a number of generalizations can be made, however.

6.7.1. Urban Environments

Urban settings are plagued with large concentrations of metallic debris and rubbish, iron and steel fencing, buried pipes, lamp posts, signage, electromagnetic fields, people with their metallic adornments and cell phones, passing and parked automobiles, and the like, all to the point where useful survey geophysical results might be impossible. On top of these factors, given that urban landscapes frequently have undergone extensive and intensive reworking, discerning pattern associated with culturally produced

anomalies can be extremely challenging and even unproductive in the complex deposits that occur in these settings.

6.7.2. Non-urban Settings

In general, remote sensing works best in open fields with uniform ground cover. Croplands must be surveyed when crops are down and years of plowing can reduce the possibility of remote detection or even eliminate the subsurface archaeology if shallow. Heavily vegetated or wooded landscapes impede movement of ground-based instruments and forests hamper visualization from the air. Steep slopes also make for difficult ground-based surveys. Even in non-urban settings landscaping such as field leveling and modern pipeline and transmission line intrusions must be watched for.

6.7.3. Aerial (and Space) Methods

The detection of archaeological features from the air is best suited to two environments. The first is arid landscapes with sparse vegetation where archaeological ruins on the surface may easily be witnessed. The second lies in farmland under uniform crop types, as nearly a century of work in European aerial archaeology has shown (Wilson 2000). In farmland, each individual plant acts as a "sensor" to conditions below; its health may be stunted by certain kinds of archaeological features (stone walls and pavements) and promoted by others (buried ditches, pits, and middens). These variations in health are reflected in plant color, height, and other factors that may be witnessed from the air. Such "crop marking" should be visible across the country, but is generally limited to robust architectural remains, and is dependent on season and stages of crop development.

6.7.4. Geophysics

Geophysical instrumentation is designed for a wide range of conditions, but each has environmental limitations.

Magnetometry. This technique responds largely to magnetic particles in the soil, but in young or undeveloped soils native magnetic susceptibility may be too low for detection purposes. This frequently occurs in the Southwest where soil development is poor and was witnessed in this project at Pueblo Escondido where magnetic anomalies were extremely subtle. Magnetometry is also not useful in volcanic areas or regions of igneous rock outcropping, owing to high levels or remanent magnetism. Iron or steel debris, when abundant, can preclude useful results with this method.

Resistivity. This method requires soil moisture, and so is limited to winter and spring surveys in much of the arid West. Even in more moderate climates sufficient soil moisture may be a problem as the landscape dries out between periods of rain. It also cannot be undertaken in winter when the ground is frozen, a circumstance that can produce near-infinite resistivity. Shallow or rising bedrock can also produce spurious readings. A benefit of resistivity is that in urban contexts it is insensitive to metallic artifacts and electromagnetic fields, allowing recovery of this data type.

EM Induction. This method largely parallels resistivity, but it may be possible to acquire data when the surface is dry because probes are not used and electromagnetic fields can be induced in subsurface deposits. A drawback of these methods is that they

are also sensitive to metals of any kind. Results can be degraded when metallic litter is abundant.

GPR. The principal shortcoming of this technique is too much moisture and clay deposits. Both promote conductivity, which disperses radar energy and limits penetrating. While it is true that water can sometimes enhance contrasts (when particular subsurface features concentrate water) in general it reduces the method's utility. Additionally, GPR is sometimes too sensitive, picking up every tree root, rodent hole, cavity, and rock in the soil and generating too many anomalies that may obfuscate detection of culturally significant ones when soils are not uniform.

Characteristics of key geophysical instruments are summarized in Table 6.1.

6.8. BEST SUBSET OF INSTRUMENTS

The best subset of instruments to employ in an archaeological project depends on a wide variety of factors. Primary considerations lie in the nature of the environment and its climate and the nature of the archaeological remains to be detected.

6.8.1. Environment

The nature of the environment must first be considered, as discussed in the previous section. In volcanic areas, magnetometry should be avoided. In arid or frozen environments, resistivity should not be considered. Instruments sensitive to electromagnetic interference should not be employed in urban settings.

6.8.2. Nature of Archaeological Remains

The second consideration in selecting instrumentation lies in the type of archaeological site being investigated, the depth to cultural features, the typical size of targets to be located, and the amount of area to be investigated.

Large areas. Aerial and space imaging allow quickest detection over large areas if the site is in a suitable context that will yield crop or vegetation marking or shadow marking if shallow. If a ground-based method is required to locate subsurface anomalies then magnetic gradiometry should be considered. It is the most rapid data collection method with a single instrument allowing 1-2 ha of survey per day. EM instruments can also allow rapid coverage of large areas.

Table 6.1. Characteristics of four principal subsurface prospecting methods.

		Magnetic	Resistivity	EM	GPR
Units:		nT	ohm/m	mS/m or ppt	ns
Common depth:		< 1.5 m	.25-2 m	.75-6 m quad 0-.5 m in-phase	*500 MHz: .5-3 m 300 MHz: 1-9 m
Typical sampling density:	**Low:**	1/m	1/m	1/m	>1 m transect spacing; 10 traces/m
	High:	16/m	4/m	4/m	.25-.5 m transect spacing; 50 traces/m
Survey time (20 m grid of 20 lines):		20 min.	45 min.	20 min.	30 min.
Area/day:		.5-2 ha	.5 ha	.5-1 ha	.25-.75 ha
Sensitivity to metals:		ferrous only	no	any	any
Situations to avoid:		metallic debris, igneous areas	surface very dry, saturated earth, shallow bedrock	high resistance areas, very dry or saturated earth, metallic debris	highly conductive clays, salts, rocky glacial deposits (e.g., moraines)
Tree effects:		impede survey, invisible in data	impede survey, positive anomaly	impede survey, negative anomaly	impede survey, roots yield anomalies
Advantages:		speed, hearths, burned areas detectable	good feature definition, moisture differences, rock; specific depth settings	speed, ease of use, collect conductivity & MS simultaneously	vertical profiles, stratigraphy, results in real time
Disadvantages:		restricted depth, need open parkland for speed, iron clutter detrimental, sensor facing critical, constant pace of movement	probe contacts slow, must deal with cables	less spatial detail, metal clutter detrimental, must maintain constant ground angle, need open parkland for speed	equipment bulky, difficult data processing, interpretations difficult
Daily data volume:		high	low	low	high
Data processing complexity:		moderate	low	moderate-low	high
Costs (USD):		$5k-25k	$600-15k	$6k-18k	$15k-30k

***** depends on soil properties**.**

Deep sites. Most geophysical instruments are designed for near-surface detection, primarily sub-meter. For survey below a meter, certain EM instruments (the EM31 by Geonics Ltd.) allow soil conductivity prospecting to a depth of 6 m, but spatial resolution is low. The same is true of electrical resistivity, which has no real depth limitation with wide probe separations, but spatial resolutions suffers with depth. If soils are not too conductive, GPR with a low frequency antenna (e.g. 200-300 MHz) can allow detection to depths of several meters.

Size of Targets. When targets to be located in an archaeological site are very small one must utilize very high sampling densities to detect contrasts and form anomalies. Small targets generally are sub-meter in size and include archaeological categories such as post holes and small pits, for example. Graves may fall in this category as well. In general, GPR is best suited for resolving small targets because from dozens to more than a hundred traces may be sent to the subsurface each linear meter, giving best chance of detecting small targets. Magnetic gradiometry offers next highest sampling densities with 10-20 measurements per meter typical.

Type of archaeology. If the site being investigated contains constructions of stone then resistivity survey is a must. It is also highly revealing when large ditch or sunken features (pits, cellars, subterranean dwellings) may be present because it is so able to detect moisture differences. Magnetometry is useful whenever magnetically enriched topsoil is reworked and redeposited. Mounded topsoil features are highly detectable, as are pits and depressions filled with topsoil. Areas of topsoil removal (e.g., ditches) also appear as negative anomalies. Any form of intensive burning enhances soil magnetism, making hearths and burned houses readily detectable. GPR responds well to most subsurface archaeological features, but is less sensitive to magnetic ones typically detected by magnetometry. EM induction instruments are largely redundant on resistivity and magnetometry. Moreover, the magnetic susceptibility component is sensitive to only very shallow targets (less than a half-meter) and experience suggests that the soil conductivity component resolves less detail than corresponding resistivity surveys.

6.8.3. Considerations of Dimensionality

Weymouth (1986) has suggested that magnetometry, resistivity, and GPR generally respond to independent dimensions of the subsurface: magnetism, conductivity, and dielectric contrasts. To this might be added thermal properties. In general, different anomalies may theoretically be indicated in each domain suggesting that multi-method surveys are warranted whenever possible (i.e., when environmental conditions allow).

6.9 COST/BENEFIT OF TECHNOLOGY

Recent events, including the conflict in Iraq and the war on terror, have underscored the fact that realistic military training is more important than ever before. As military vehicle and weapons systems evolve, there is an ever increasing need for large contiguous land areas for effective training. Archaeological sites categorized as eligible or potentially eligible for the NRHP currently are normally avoided, fragmenting the lands available for training. One way to reduce this problem is to mitigate sites located within key training areas. Using archaeo-geophysical techniques to identify sites and portions of sites that include important subsurface deposits could dramatically reduce the

costs associated with site mitigation. These amounts can be estimated, although actual percentages will depend on the training needs and nature of the archaeological resource base at individual installations, as well as other factors. The Army presently manages ca. 90,000 archaeological sites (source: Army Environmental Center). We can assume that a majority of these (e.g., 66%) are relatively small sites that lack complex cultural deposits. The remaining sites (n=30,000) are relatively large and complex. Only a fraction of these sites--for present purposes, we can assume 10% (n=3,000)--are located in areas critically needed for military training. A carefully conceived CRM plan might propose mitigation of 20% of these sites (n=600), with the rest made available for military training. The traditional approach to mitigation involving excavation of a large portion of each site is very costly. For present purposes, we estimate the cost of traditional mitigation at $100,000 per site. The total cost to mitigate 600 sites is therefore $60,000,000. If the methods used in this project were in wide use, we believe that the mitigation cost per site (and total cost for 600 sites) could be reduced by at least 50%, representing a cost avoidance of $30,000,000. If the 600 sites were mitigated over the course of 10 years, the cost per year would be $6,000,000 for traditional approaches vs. $3,000,000 for a strategy based on archaeo-geophysics and targeted excavation. The cost savings for the latter would be $3,000,000 per year, or $15,000,000 over five years. Note that these estimates are not adjusted for inflation.

6.10 TRANSITION STATUS

As the previous section illustrates, it is not unreasonable to anticipate that there could be dramatic direct cost savings to DoD if the methods applied in this project were widely used. In addition to these direct savings we believe, as many of the sections in this chapter document, that the quality of information will increase and the likelihood of unintended damage to resources and the unintended exposure of Native American human remains will be reduced. We believe that there are three major impediments to the rapid adoption of this approach. These are (1) the absence of a pool of qualified archae-geophysical practitioners, (2) the extremely time consuming and complex software processing that is currently the "state-of-the-art" and (3) the lack of awareness and acceptance of the methods by the regulatory agencies (e.g. SHPOs, Tribal Historical Preservation Offices (THPO) and Advisory Council). Issues 2 and 3 are currently being directly addressed by a recently initiated ESTCP Project: *Streamlined Archaeo-Geophysical Data Processing and Integration for DoD Field Use*

The ESTCP project has two primary objectives: 1) assemble a single software package,
ArchaeoMapper, that will serve as an effective medium for infusing the integrated, multi-sensor geophysical approach into wide use; and 2) demonstrate and validate the cost and performance benefits of the approach and technology infusion tool, in conjunction with the annual NPS *Current Archeological Prospection Advances for Non-destructive Investigations in the 21st Century Workshop.* The multi-agency demonstration will involve DoD geophysicists, representatives of federal, state, and THPOs, CRM practitioners, and federal and state resource managers. The project should also serve to address aspects of the first impediment by providing an easier-to-use software environment that should serve to increase the number of practitioners. In addition, as the value of archaeo-geophysical methods becomes increasingly obvious from projects such as this and many others, the number of interested archaeologists in growing.

7. REFERENCES CITED

Ahler, S. R., D. L. Asch, D. E. Harn, B. W. Styles, K. White, C. Diaz-Granados, and D. Ryckman. 1999. *National Register Eligibility Assessments of Seven Prehistoric Archaeological Sites at Fort Leonard Wood, Missouri.* Submitted to U.S. Army Construction Engineering Research Laboratory, Champaign, Illinois, by Illinois State Museum Society, Quaternary Studies Program, Technical Report 98-1202-28, Springfield.

Ahler, Stanley A., and Kenneth L. Kvamme. 2000. New Geophysical And Archaeological Investigations At Huff Village State Historic Site (32MO11), Morton County, North Dakota. Submitted to the State Historical Society of North Dakota, Bismarck.

Avery, T.E., and Berlin G.L. 1992. *Fundamentals of Remote Sensing and Airphoto Interpretation, 5th Ed*. Macmillan, New York.

Anyon, Roger. 1985. Archaeological Testing at the Fairchild Site (LA 45732), Otero County, New Mexico. Office of Contract Archaeology, University of New Mexico, Albuquerque.

Bales, Jenny R., and Kenneth L. Kvamme. 2004.Geophysical Signatures Of Earthlodges In The Dakotas. In *Plains Earthlodges: Perspectives On A Major North American House Type*. Donna Roper and Elizabeth Pauls, (Eds.), Tuscaloosa: University Of Alabama Press.

Becker, H. 1995. From Nanotesla to Picotesla - A New Window for Magnetic Prospecting in Archaeology. *Archaeological Prospection* 2:217-228.

Beckes, Michael R. 1975. U.T. Austin-Site Survey Form for Site M-410, FB9569 (Escondida Pueblo).

Bevan, B.W. 1998. *Geophysical Exploration for Archaeology: An Introduction to GeophysicalExploration*. Midwest Archeological Center Special Report 1. Lincoln, Nebraska Bewley, Robert H. (2000. Aerial photography for archaeology, In *Archaeological Method And Theory: An Encyclopedia*: Linda Ellis (ed.), pp. 3-10. New York: Garland Publishing.

Braun, Lucy E. 1950. *Deciduous Forests of Eastern North America*. MacMillan, New York.

Brizzolari, E., Ermolli, F., Orlando, L., Piro, S., Versino, L., 1992, Integrated geophysical methods in archaeological surveys. *Journal of Applied Geophysics* 29:47-55.

Burrough, P.A., and McDonnell, R.A. 1998. *Principles of Geographical Information Systems*. Oxford University Press, Oxford.

Buteux, S., Gaffney, V., White, R., and van Leusen, M. 2000. Wroxeter Hinterland Project and geophysical survey at Wroxeter. *Archaeological Prospection* 7:69-80.

Canny, John F. 1986. A Computational Approach to Edge Detection. IEEE Transaction on Pattern Analysis and Machine Intelligence. 8(6): 679-698.

Chavez, P. S. Jr., S. C. Sides, and J. A. Anderson. 1991. Comparison of Three Different Methods
to Merge Multiresolution and Multispectral Data: Landsat TM and SPOT Panchromatic.
Photogrammetric Engineering and Remote Sensing 57(3): 295-303.

Clark, A. 2000. *Seeing Beneath the Soil: Prospection Methods in Archaeology.* Routledge, London.

Clark, W.A., and P.L. Hosking. 1986. *Statistical Methods for Geographers (Chapter 13)*. John Wiley & Sons, New York Clay, R.B., 2001, Complementary geophysical survey techniques: why two ways are always better than one. *Southeastern Archaeology* 20:31-43.

Conyers, L.B. 2004. *Ground-penetrating Radar for Archaeology*. AltaMira Press, Walnut Creek, California.

Conyers, Lawrence B. 2006. Ground-Penetrating Radar. In *Remote Sensing in Archaeology: An Explicitly North American Perspective*, Jay K. Johnson (ed.), pp. 131-160. University of Alabama Press.

Cottier, John W. 1977. *Lawson Field: A Cultural Resource Survey and Evaluation of a Selected Portion of Fort Benning Military Reservation*. Department of Sociology and Anthropology, Auburn University, Alabama. Report submitted to Fort Benning.

Crawford, O.G.S. and A. Keiller. 1928. *Wessex from the Air.* Oxford: Clarendon Press.

Ciminale, M. & M. Loddo. 2001. Aspects of magnetic data processing, *Archaeological Prospection* 8:239-246

Dabas, M., and A. Tabbagh. 2000. Magnetic prospecting. In *Archaeological Method and Theory: An Encyclopedia*, Linda Ellis (ed.), Garland Publishing, New York.

David, A. 2001, Overview—the role and practice of archaeological prospection. In *Handbook Of Archaeological Sciences*, D.R. Brothwell and A.M. Pollard (eds.), pp. 521-527. John Wiley, New York

Davis, J.C. 2002. *Statistics and Data Analysis in Geology, 3rd Ed*. John Wiley, New York Deuel, Leo. 1969. Flights into Yesterday. St. Martin's Press, New York.

Doneus, M., and W. Neubauer. 1998. 2D combination of prospection data. *Archaeological Prospection* 5:29-56.

Drager, Dwight L. & Thomas R. Lyons. 1985. *Remote sensing photogrammetry in archeology: the Chaco mapping project*. Albuquerque, New Mexico: National Park Service.

Fassbinder, J., H. Stanjek & H. Vali. 1990. Occurrence of magnetic bacteria in soil, *Nature* 343:161-163.

Foster, H. Thomas II. 2005a. (in review) *The Archaeology of the Lower Creek Indians, 1715-1836*. University of Alabama Press.

Foster, H. Thomas II. 2005b. *Excavations at the Muskogee Town of Cussetuh (9CE1)*. Draft report submitted to the Engineer Research and Development Center, Construction Engineering Research Laboratory, Champaign, IL, by BHE Environmental, Inc. Cincinnati, OH.

Fowler, Martin J.F. 1996. High-resolution satellite imagery in archaeological application: a Russian satellite photograph of the Stonehenge region, *Antiquity* 70:667-671.

Gaffney, C., and Gater, J. 2003, *Revealing the Buried Past: Geophysics for Archaeologists*. Tempus Publishing, Stroud, England.

Gaffney, C., Gater, J.A., Linford, P., Gaffney, V., and White, R. 2000. Large-scale systematic fluxgate gradiometry at the Roman city of Wroxeter. *Archaeological Prospection* 7:81-99.

Gaffney, V.L., and P.M. van Leusen. 1996. Extending GIS methods for regional aarchaeology: the Wroxeter Hinterland Project. In Kammermans, H., and Fennema,

K. (eds.) *Interfacing the Past: Computer Applications and Quantitative Methods in Archaeology, 1995*, pp. 297-305, Analecta Praehistorica Leidensia 28, Leiden University Press, Leiden.

Geoarchaeology Research Associates. 2000. *Intensive Survey of Black Ramp Extension Project Lawson Army Airfield, Fort Benning, Georgia: Preliminary Summary*. Report submitted by Geoarchaeology Research Associates, 5912 Spencer Avenue, Riverdale, NY, to Panamerican Consultants, Inc., 924 26th Avenue East, Tuscaloosa, AL. December 4, 2000.

Geoscan Research. 2000. *Geoplot Version 3.00 for Windows, Instruction Manual.* Geoscan Research, Bradford, England.

Gonzales, J. E. 2005. Urban Heat Islands Developing in Coastal Tropical Cities. EOS, Transactions, American Geophysical Union. Volume 86, Number 42, 18 October.

Goodman, D., Nishimura, Y. & J.D. Rogers (1995. GPR time-slices in archaeological prospection, *Archaeological Prospection* 2:85-89.

Hailey, T.I. 2005. The powered parachute as an archaeological aerial reconnaissance vehicle. *Archaeological Prospection* 12:69-78.

Hargrave, Michael L. 1999a. *A Comparison of Traditional and Geophysical Strategies for Assessing the National Register Status of Archaeological Sites at Fort Riley,Kansas*. Special Report 99/22/January 1999. U. S. Army Construction Engineering Research Laboratory, Champaign, IL.

Hargrave, Michael L. 1999b. *Geophysical and Archaeological Investigations of Historic Sites at Fort Riley, Kansas*, by Thomas K. Larson, Lewis E. Somers, Dori M. Penny, and Michael L. Hargrave. Technical Report 99/47/ June 1999. U. S. Army Construction Engineering Research Laboratory, Champaign, IL.

Hargrave, Michael L. 2006. Ground truthing the results of geophysical surveys. In *Geophysical and Airborne Remote Sensing Applications in Archaeology: A Guide for Cultural Resource Managers*, J. Johnson (ed.), pp. 269-304, University of Alabama Press, Tuscaloosa.

Hargrave, Michael L., Charles R. McGimsey, Mark J. Wagner, Lee A. Newsom, Laura Ruggiero, Emanuel Breitburg, and Lynette Norr. 1998. *The Yuchi Town Site (1RU63), Russell County, Alabama: An Assessment of the Impacts of Looting*. Report submitted to the Cultural Resources Management Program at Ft. Benning by the Cultural Resources Research Center, U.S. Army Construction Engineering Research Laboratory. USACERL Special Report 98/48. February, 1998.

Hargrave, M.L., Somers, L.E., Larson, T.K., Shields, R., and Dendy, J. 2002. The role of resistivity survey in historic site assessment and management: an example from Fort Riley, Kansas. *Historical Archaeology* 36: 89-110.

Hedrick, Mrs. John A. 1967. Escondida Survey. *The Artifact* 5(2):19-24.

Herron, Tammy Forehand, and Robert Moon. 2005. G*round Truthing of a Multi-Sensor Remote Sensing Survey at the George Galphin Site, Silver Bluff Audubon Sanctuary, Aiken County, South Carolina.* Draft report submitted to the Engineer Research and Development Center, Construction Engineering Research Laboratory, Champaign, IL, by the Savannah River Archaeological Research Program, South Carolina Institute of Archaeology and Anthropology, University of South Carolina, Columbia.

Hesse, Albert. 2000. Archaeological prospection, In *Archaeological Method And Theory: An Encyclopedia,* Linda Ellis (ed.), New York: Garland Publishing.

Hinton, J.C. 1996. GIS and remote sensing integration for environmental applications. *International Journal of Geographical Information Systems* 10:877-890.

Hodgson, M. E., J. R. Jensen, J. A. Tullis, K. D. Riordan, And C. M. Archer. 2003. Synergistic Use of LIDAR and Color Aerial Photography for Mapping Urban Parcel Imperviousness. *Photogrammetric Engineering and Remote Sensing* 69: 973-980.

Hosmer, D.W., and Lemeshow, S. 2000. *Applied Logistic Regression, 2nd ed*. John Wiley, New York.

Hudson, Charles. 1976. *The Southeastern Indians*. The University of Tennessee Press.

Jantz, Donald R., Rodney F. Horner, Harold T. Rowland, and Donald A. Gier (1975. *Soil Survey of Riley County and Part of Geary County, Kansas*. Bulletin No. 39. Geological Survey, University of Kansas, Lawrence.

Jensen, J. R. 1996. Introductory Digital Image Processing: A Remote Sensing Perspective, 2nd ed. Prentice Hall, 316 p

Jensen, J. R. 2000. Remote Sensing of the Environment: An Earth Resource Perspective. Prentice Hall, Upper Saddle River, New Jersey.

Jensen, J. R. 2005. *Introductory Digital Image Processing: A Remote Sensing Perspecitve*. Prentice Hall, Upper Saddle River, New Jersey.

Johnson, J.K. and Haley, B.S. 2004. Multiple sensor applications in archaeological geophysics. In *Proceedings of SPIE Vol. 5234, Sensors, Systems and Next Generation Satellites VII*, R. Meynart, S.P. Neeck, H. Simoda, J.B. Lurie and M.L. Aten (eds.), pp. 688-697. SPIE, Bellingham, Washington.

Johnson, Jay K. and Haley B.S.. 2006. A cost-benefit analysis of remote sensing application in cultural resource management archaeology. In *Geophysical and Airborne Remote Sensing Applications in Archaeology: A Guide for Cultural Resource Managers*, J. Johnson, ed., pp. 33-45, University of Alabama Press, Tuscaloosa.

Jurney, D. 2001. *The Effectiveness of Survey Techniques in the Ozarks*. Paper presented at the annual meeting of the Arkansas Archeological Society, Hot Springs.

Kennedy, David. 1998. Declassified satellite photographs and archaeology in the Middle East: case studies from Turkey, *Antiquity* 72:553-561.

Klecka, W.R. 1980. *Discriminant Analysis*. Sage University Paper Series on Quantitative Applications in the Social Sciences, 07-019. Sage Publications, Beverly Hills, CA.

Krakker, J.J., M.J. Shott, and P.D. Welch. 1983. Design and evaluation of shovel-test sampling in regional archaeological survey. *Journal of Field Archaeology* 10:469-480.

Kreisa, Paul P., and Gregory R. Walz. 1997. *Archaeological Test Excavations of Four Sites at Fort Riley, Riley and Geary Counties, Kansas*. Public Service Archaeology Program Research Report No. 29. Report submitted to the U. S. Army Construction Engineering Research Laboratory, Champaign, IL, by the University of Illinois at Urbana-Champaign.

Kresja, Paul P., and Jacqueline M. McDowell. 2005. *Archaeological Ground Truthing of Remote-Sensing-Derived Anomalies at Army City, Fort Riley, Kansas*. Public Service Archaeology Program Research Report No. 87. Draft report submitted to the Engineer Research and Development Center, Construction Engineering Research Laboratory, Champaign, IL, by the University of Illinois, Urbana-Champaign.

Kroeber, A. L. 1963. *Cultural and Natural Areas of Native North America*. University of California Press.

Kvamme, K.L. 1989. Geographic information systems in regional archaeological research and data management. *Archaeological Method and Theory*, *Vol. 1*, M.B. Schiffer (ed.), pp. 139-202, University of Arizona Press, Tucson.

Kvamme, K.L. 2001. Archaeological prospection in fortified Great Plains villages: new insights through data fusion, visualization, and testing. In *Archaeological Prospection: 4th International Conference on Archaeological Prospection*, P.M. Doneus, A. Eder-Hinterleitner, W. Neubauer (eds.), pp. 141-143. Austrian Academy of Sciences Press, Vienna.

Kvamme, K.L. 2003. Geophysical surveys as landscape archaeology. *American Antiquity* 68:435-457.

Kvamme, K. L. 2005. There and Back Again: Revisiting Archaeological Location Modeling, In *GIS and Archaeological Predictive Modeling*, M.W. Mehrer and K. Wescott, eds., CRC Press, New York.

Kvamme, K. L. 2005a. Integrating Multidimensional Geophysical Data. *Archaeological Prospection*.

Kvamme, K. L. 2005b. Magnetometry: Nature's Gift to Archaeology. In Jay K. Johnson (ed.), *Geophysical and Airborne Remote Sensing Applications in Archaeology: A Guide for Cultural Resource Managers*. Tuscaloosa: University of Alabama Press.

Kvamme, K.L. 2006. Remote sensing: archaeological reasoning through physical principles and pattern recognition. In *Archaeological Concepts for the Study of the Cultural Past*, A.P. Sullivan III (ed.). University of Utah Press, Salt Lake City.

Kvamme, K.L. 2006. Integrating Multidimensional Geophysical Data. *Archaeological Prospection.*

Kvamme, J.C., T.I. Hailey, and K.L. Kvamme. 2004. Aerial Archaeology at Double Ditch Site
Historic Site, North Dakota. Paper presented at the 62nd Plains Conference, October 13-16, Billings, Montana.

Larson, Thomas K., and Dori M. Penny. 1998. Results of Archaeological Ground Truthing Investigations at Historic Sites, Fort Riley, Kansas. Report submitted to the U. S. Army Construction Engineering Research Laboratory, Champaign, IL, by LTA, Inc., Laramie, WY.

Leckebusch, J. 2003. Ground-penetrating radar: a modern three-dimensional prospection method. *Archaeological Prospection* 10: 213-240.

Lehmer, Donald J. 1948. The Jornado Branch of the Mogollon, University of Arizona *Social Science Bulletin* No. 17, Tucson.

Levin, S. A. 1992. The Problem of Pattern and Scale in Ecology. *Ecology* 73: 1943-1967.

Lillesand, T.M., and R.W. Kiefer. 1994. *Remote Sensing and Image Interpretation*. John Wiley, New York.

Lockhart, J.J., and Thomas J. Green. 2006. The current and potential role of archaeogeophysics in cultural resource management in the United States. In *Geophysical and Airborne Remote Sensing Applications in Archaeology: A Guide for Cultural Resource Managers*, J. Johnson, ed., pp. 17-32, University of Alabama Press, Tuscaloosa.

Luvall, J. C., D. Rickman and M. Estes. 2005. Aircraft Based Remotely Sensed Albedo and Surface Temperatures for Three US Cities. Paper presented at the Cool Roofing: Cutting Through the Glare Roofing Symposium, Roof Consult. Inst. Found., Atlanta, GA.

Mauldin, Raymond. 1986. Settlement and Subsistence Patterns During the Pueblo Period on Fort Bliss, Texas: a Model. In *Mogollon Variability*, C. Benson and S. Upham (eds). The University Museum Occasional Papers, No. 15, pp. 225-270. New Mexico State University, Las Cruces.

Miller, Myles R., and Nancy A. Kenmotsu. 2004. Prehistory of the Jornada Mogollon and Eastern Trans-Pecos Regions of West Texas. In *The Prehistory of Texas*, T. K. Perttula (ed), pp. 205-265. Anthropology Series No. 9. Texas A&M University Press, College Station.

Moik, J.G. 1980. *Digital Processing of Remotely Sensed Images*. National Aeronautics and Space Administrations, Washington, D.C.

Muchoney, D., S. Gopal, J. Hodges, N. Morrow, A. Strahler, J. Borak, H. Chi, And M. Friedl. 2000. Application of the MODIS Global Supervised Classification Model to Vegetation and Land Cover Mapping of Central America. *International Journal of Remote Sensing* 21: 1115-1138.

Mussett, Alan E. & M. Aftab Khan. 2000. *Looking into the Earth: an introduction to geological geophysics*. Cambridge: Cambridge University Press.

Neubauer, W., Eder-Hinterleitner, A. 1997. Resistivity and magnetics of the Roman town Carnuntum, Austria: an example of combined interpretation of Prospection Data. *Archaeological Prospection* 4:179-189.

O'Laughlin, Thomas C. 2001. Long Lessons and Big Surprises: Firecracker Pueblo. In *Following Through: Papers in Honor of Phyllis S. Davis*, R. Wiseman, T. C. O'Lauhglin and C. T. Snow (eds), pp. 115-131. Archaeological Society of New Mexico 27, Albuquerque.

O'Laughlin, T.C., and D.L. Martin. 1993. Phase II Additional Testing Loop 375 Archaeological Project, El Paso County, Texas. *The Artifact* 31(4).

O'Sullivan, D., and D. J. Unwin. 2003. Geographic Information Analysis. John Wiley & Sons, Inc., Hoboken, New Jersey.

O'Steen, Lisa D., John S. Cable, Mary Beth Reed, and J. W. Joseph. 1997. *CulturalResources Survey, Lawson Army Airfield, Ft. Benning, Georgia and Alabama.* Report submitted to the National Park Service, Southeast Archaeological Center, Tallahassee, FL, by New South Associates, Inc., Stone Mountain, GA.

Openshaw, S., and C. Openshaw. 1997. Artificial Intelligence in Geography. John Wiley & Sons, Chichester, U.K.

Piro, S., Mauriello, P., and Cammarano, F. 2000. Quantitative integration of geophysical methods for archaeological prospection. *Archaeological Prospection* 7: 203-213.

Proudfoot, B. 1976. The Analysis and Interpretation of Soil Phosphorous in Archeological Contexts. Geoarchaeology. Westview Press, Boulder, CO.

Raber, G. T., J. A. Tullis, and J. R. Jensen. 2004. LIDAR Statistical Image Fusion with IKONOS Data for Land Cover Classification. Association of American Geographers 100th Annual Meeting, Philadelphia, Pennsylvania.

Reeves, Dache. 1936. Aerial photography and archaeology, *American Antiquity*.
Renfrew, Colin, and Paul Bahn. 2000. *Archaeology, 3rd edition*. Thames and Hudson, New York.
Rion, George P. 1960. *Army City, Kansas: The History of a World War I Camptown*. Master's thesis, Department of History, Political Science, and Philosophy, Kansas State University of Agriculture and Applied Science, Manhattan, KS.
Rogers, Vergil A. 1985. *Soil Survey of Aiken County Area, South Carolina*. United States Department of Agriculture, Washington, D.C.
Russell, S. J., and P. Norvig. 2003. Artificial Intelligence: A Modern Approach. Prentice Hall, Upper Saddle River.
Sabins, Floyd F. 1997. *Remote sensing: principles and interpretation, 3rd edition*. New York: W.H. Freeman.
Schiffer, Michael B. 1976. *Behavioral Archaeology*. New York: Academic Press.
Schott, J.R. 1997 Remote Sensing: The Image Chain Approach. Oxford University Press.
Schmidt, Armin. 2001. *Geophysical data in archaeology: a guide to good practice*. Oxford: Oxbow Books.
Schowengerdt, R.A. 1983. *Techniques for Image Processing and Classification in Remote Sensing*. Academic Press, New York.
Schowengerdt, R.A. 1997. *Remote Sensing: Models and Methods for Image Processing*. Academic Press, San Diego.
Scollar, I., Tabbagh, A., Hesse, A., and Herzog, I. 1990. *Archaeological Prospection and Remote Sensing*. Cambridge University Press, Cambridge.
Scurry, James D., J. Walter Joseph, and Fritz Hamer. 1980. *Initial Archaeological Investigations at Silver Bluff Plantation, Aiken County, South Carolina.* Research Manuscript Series. Institute of Archaeology and Anthropology, University of South Carolina, Columbia.
Seaman, Timothy J., William H. Doloman, and Richard C. Chapman (eds). 1988. *The Border Star 85 Survey, Toward an Archaeology of Landscapes*. Report submitted to the U.S. Army Corps of Engineers, Ft. Worth District, by the Office of Contract Archaeology, University of New Mexico.
Sever, T.L. 2000. Remote Sensing Methods. Chapter 2, in *Science and Technology in Historic*
Preservation, Ray Williamson and Paul Nickens (eds). Advances in Archaeological and Museum Science, volume 4. Kluwer Academic and Plenum Press.
Sever, T.L. 1998. Archaeology. NASA.
Sever, T.L. 1998. Appendix I: Wright-Patterson Remote Sensing Report." In Archaeological, Geophysical, and Remote Sensing Investigation of the 1910 Wright Brother's Hanger, Wright- Patterson Air Force Base, Ohio, by David Babson, Michael L. Hargrave, Thomas L. Sever, John S. Isaacson, and James A. Zeidler. Cultural Resources Research Center, U.S. Army Construction Engineering Research Laboratories, Champaign, IL.
Sever, T.L. 1990. Remote Sensing Applications in Archeological Research: Tracing Prehistoric Human Impact Upon the Environment. Doctoral dissertation, University of Colorado. University Microfilms, Ann Arbor, Michigan. 1990.

Sever, T.L. 1983. Feasibility Study to Determine the Utility of Advanced Remote Sensing Technology in Archeological Investigations. Report No. 227. NASA, Stennis Space Center, Science and Technology Laboratory, SSC, MS. December, 1983.

Shelford, Victor E. 1963. *The Ecology of North America*. University of Illinois Press, Urbana, IL.

Shott, M. 1985. Shovel-test sampling as a site discovery technique: a case study from Michigan. *Journal of Field Archaeology* 12: 457-468.

Somers, Lewis E. 1997. *Geophysical Investigations at Archaeological Sites 14RY3183; 14RY3193; 14RY5155, and 14GE3183, Fort Riley, Kansas*. Report submitted to the U.S. Army Construction Engineering Research Laboratory, Champaign, IL, by Geoscan Research (USA), Sea Ranch, CA.

Somers, Lewis E. 1998. *Geophysical Survey of Historic Sites at Fort Riley, Kansas: Farmsteads 14GE1108, 14RY152, 14RY2118, 14RY2170, and14RY2171, and the Army City Site,14RY3193*. Report submitted to the U. S. Army Construction Engineering Research Laboratory, Champaign, IL, from Geoscan Research (USA), Sea Ranch, CA.

Sullivan, Alan P. III (ed.). 1998. *Surface Archaeology*. University of New Mexico Press, Albuquerque.

Swan, Caleb. 1855. Position and State of Manners and Arts in the Creek, or Muscogee Nation in 1791. In *Information Respecting the History, Condition and Prospects of the Indian Tribes of the United States*, Henry Rowe Schoolcraft (ed.), 5:251-283. J. B. Lippincott & Company, Philadelphia

Swanton, John R. 1979. The Indians of the Southeastern United States. Smithsonian Institution Press, Washington, D.C.

Telford, W.M., Geldart, L.P., Sheriff, R.E. 1990. *Applied Geophysics, 2nd ed*. Cambridge University Press, Cambridge

Toom, D. L., and K. L. Kvamme. 2002. The "Big House" At Whistling Elk Village (39HU242): Geophysical Findings And Archaeological Truths. *Plains Anthropologist* 47:5-16.

TRC Environmental. 2005. *Ground Truthing Remote Sensing Data at the Escondida Site (LA458), Otero County, New Mexico*. Draft report submitted to the Engineer Research and Development Center, Construction Engineering Research Laboratory, Champaign, IL, by TRC Environmental, El Paso, TX.

Trucco, Emanuele and Alessandro Verri. 1998. Introductory Techniques for 3-D Computer Vision. Upper Saddle River, NJ.: Prentice Hall, Inc

Tullis, J. A. 2003. Data Mining to Identify Optimal Spatial Aggregation Scales and Input Features: Digital Image Classification with Topographic Lidar and Lidar Intensity Returns. Doctoral Dissertation, University of South Carolina, Columbia, SC.

Tullis, J. A., and J. R. Jensen. 2003. Expert System House Detection Using Size, Shape, and Context. *Geocarto International* 18: 5-15.

U.S. Army Engineer Topographic Laboratories. 1976. Fort Benning, Georgia, Terrain Analysis. The Terrain Analysis Center, U.S. Army Engineer Topographic Analysis Center, U.S. Army Engineer Topographic Laboratories, Fort Belvoir, VA.

van der Sande, C., de Jong, S.M., and de Roo, A.P.J. 2003. A Segmentation and Classification Approach to IKONOS-2 Imagery for Land Cover Mapping to Assist

Flood Risk and Flood Damage Assessment. *International Journal of Applied Earth Observation and Geoinformation* 4:217-229.

van Leusen, M. 2001. Archaeological data integration. In *Handbook Of Archaeological Sciences*, D.R. Brothwell and A.M. Pollard (eds.), pp. 575-583. John Wiley, New York.

Waselkov, Gregory A., and Kathryn E. Holland Braund, (eds). 1995. *William Bartram on the Southeastern Indians*. University of Nebraska Press, Lincoln.

Weymouth, J.W. 1986. Geophysical methods of archaeological site surveying. In *Advances in Archaeological Method and Theory, Vol. 9*, M.B. Schiffer (ed.), pp. 311-395. Academic Press, New York.

Whalen, Michael E. 1994. *Turquoise Ridge and Late Prehistoric residential Mobility in the Desert Mogollon Region*. University of Utah Anthropological Papers Number 118. University of Utah Press, Salt Lake City.

Wescott, K.L., And R.J. Brandon, (eds.). 2000. *Practical Applications Of GIS For Archaeologists: A Predictive Modeling Toolkit*. London: Taylor & Francis.

Whimster, Rowan. 1989. *The emerging past: air photography and the buried landscape*. London: Royal Commission on the Historic Monuments of England.

Willey, Gordon. 1938. *Excavations at Lawson Field Site, Fort Benning Reservation, Columbus GA*. In Kasihta Town Report Records, Folders 9 and 10, National Park Service, Tallahassee.

Willey, Gordon R., and William H. Sears. 1952. The Kasita Site. *Southern Indian Studies* 4:3-18.

Wilson, D.R. 2000. *Air Photo Interpretation for Archaeologists*. Arcadia Publishing, Charleston, South Carolina.

Wynn, J.C. 1986. A Review Of Geophysical Methods Used In Archaeology. *Geoarchaeology* 1:245-257.

APPENDIX A: SUPPORTING DATA

Approximately 8 gigabytes of project data and metadata (readme) files are available via anonymous FTP from the Center for Advanced Spatial Technologies.

URL: **ftp://serdp.cast.uark.edu/serdp/**

Please direct any questions to debbie@cast.uark.edu.

APPENDIX B: LIST OF TECHNICAL PUBLICATIONS

Articles Published In Peer-Reviewed Journals

- Kvamme, Kenneth L. (2005). Integrating Multidimensional Geophysical Data. *Archaeological Prospection*. http://dx.doi.org/10.1002/arp.268 (in-print version slated for February, 2006 issue)

Technical Reports

- Foster II, H. Thomas (2005). *Excavations at the Muskogee Town of Cussetuh (9CE1), Draft Report*. BHE Environmental, Inc., report submitted to the University of Arkansas and Engineer Research and Development Center, Construction Engineering Research Laboratory, Champaign, Illinois.
- Herron, Tammy Forehand, and Robert Moon (2005). *Ground Truthing of a Multi-Sensor Remote Sensing Survey at the George Galphin Site, Silver Bluff Audubon Sanctuary, Aiken County, South Carolina.* Draft report submitted to the Engineer Research and Development Center, Construction Engineering Research Laboratory, Champaign, Illinois, by the Savannah River Archaeological Research Program, South Carolina Institute of Archaeology and Anthropology, University of South Carolina, Columbia.
- Koons, Michele L. (2005). *Visualizing Ground-Penetrating Radar Data in Three-dimensions*. University of Pennsylvania report submitted to University of Arkansas.
- Krejsa, Paul, and Jacqueline M. McDowell (2005). *Draft Archaeological Ground Truthing of Remote Sensing-Derived Anomalies at Army City, Fort Riley, Kansas*. Public Service Archaeology Program Research Report No. 87, submitted to Engineer Research and Development Center, Construction Engineering Research Laboratory, Champaign, Illinois.
- Maki, David (2003). *Ground Based Geophysical Investigations of Two Archaeological Sites.* Archaeo-Physics Report of Investigations No. 66, submitted to University of Arkansas.
- Lukowski, Paul, and Elia Perez (2005). *Ground Truthing Remote Sensing Data at the Escondida Site (LA 458), Otero County, New Mexico*. TRC Environmental report submitted to Engineer Research and Development Center, Construction Engineering Research Laboratory, Champaign, Illinois.
- Sever, Thomas L. , and Burgess F. Howell (2005). *Aerial Multi-Sensor Remote Sensing Investigation for Archeological Resource Inventory.* Report submitted to the Center for Advanced Spatial Technologies, University of Arkansas.

Conference/Symposium Proceedings

- Kvamme, Kenneth L. (2005). Archaeological Modeling with GIS at Scales Large and Small. In *Reading Historical Spatial Information from Around the World: Studies of Culture and Civilization Based on Geographic Information Systems Data.* 24th International Research Symposium, International Research Center for Japanese Studies, Kyoto, Japan, pp. 169-187.

Published Technical Abstracts

- Kvamme, Kenneth L. (2005). *Integrating Geophysics with GIS*. Indo-US Science and Technology Forum Workshop on Digital Archaeology: A New Paradigm for Visualizing the Past through Computing and Information Technology, Mussoorie, Uttaranchal, India.
- Ernenwein, E.G., and K.L. Kvamme (2005). *Geophysical Data Fusion: Combining Sensor Outputs through Graphical, Mathematical, and Statistical Approaches*. Computer Applications and Quantitative Methods in Archaeology Conference (CAA), Tomar, Portugal.
- Kvamme, K.L., and E.G. Ernenwein (2005). *Multidimensional Fusion of Geophysical and Other Data from Army City, Kansas, and Pueblo Escondido, New Mexico*. Poster presentation, Annual meeting of the Society for American Archaeology, Salt Lake City.
- Hargrave, M.L., K.L. Kvamme, and E.G. Ernenwein (2005). *Methodological Issues in Ground Truthing the Results of Remote Sensing Surveys*. Poster presentation, Annual meeting of the Society for American Archaeology, Salt Lake City.
- Kvamme, Kenneth L. (2005). *Integrating Remote Sensing Data at Army City, Kansas*, Poster presentation, Annual Conference on Historical and Underwater Archaeology, Society for Historical Archaeology, York, England.
- Limp, W.F., K.L. Kvamme, E.G. Ernenwein, D.L. Harmon, M.L. Hargrave (2005). *Fusion of Geophysical Data for Subsurface Archaeological Detection and Mapping*. Paper presented at the Partners in Environmental Technology Symposium, Washington, D.C. (2005).
- Kvamme, K.L., E.G. Ernenwein, T.L. Sever, D.L. Harmon, W. F.Limp, and M. Hargrave (2005). *Data Fusion of Archaeological Remote Sensing from Ground-, Air-, and Space-based Platforms*. Poster presented at the Arkansas GIS User Forum Conference, Hot Springs, AR.
- Kvamme, Kenneth L. (2004). *Multidimensional Remote Sensing & Data Fusion in North American Archaeological Sites*. Paper presented at the meeting of the International Union of Pre- and Proto-historic Sciences, Commission IV: Data Management and Mathematical Methods in Archaeology, Santa Fe, New Mexico.
- Ernenwein, Eileen G., and Kenneth L. Kvamme (2004). *Ground-penetrating Radar at the Landscape Scale: New Problems and Possible Solutions*. Paper presented at the Archaeological Sciences of the Americas Symposium, University of Arizona, Tucson.
- Ernenwein E.G., and K.L. Kvamme (2004). Archeo-Geophysical, Panchromatic, and Multispectral Data Synergy at Army City, Kansas. Poster presented at the Annual Conference on Historical and Underwater Archaeology, Society for Historical Archaeology, St. Louis, Missouri.
- Kvamme, Kenneth L. (2004). *Data Fusion of Archaeological Remote Sensing from Ground-, Air-, and Space-based Platforms*. Paper presented at the Partners in Environmental Technology Symposium, Washington, D.C. (2004).
- Ernenwein, Eileen G., Kenneth L. Kvamme, W. Fredrick Limp, Deborah L. Harmon, Michael L. Hargrave, Thomas L. Sever (2004). *Archaeo-Geophysical,*

Panchromatic, and Multispectral Data Synergy at Four DoD and DoE Archaeological Sites. U.S. Department of Defense Conservation Conference, Savannah, Georgia.

- Kvamme, K.L., E.G. Ernenwein, T.L. Sever, D.L. Harmon, W. F.Limp, M. Hargrave, and L.E. Somers (2003). *Archeo-Geophysical, Panchromatic, Thermal, and Multispectral Data Synergy at Four DoD And DoE Archaeological Sites*. Poster presented at the Partners in Environmental Technology Technical Symposium & Workshop, Department of Defense, Washington, D.C.
- Ernenwein, E.G., K.L. Kvamme, W.F. Limp, D.L. Harmon, M. Hargrave, T.L. Sever, and L.E. Somers (2003). *Multi-Dimensional Remote Sensing: Fusing Ground, Air, and Satellite Data from Archaeological Sites*. Poster presented at the annual meeting of the Society for American Archaeology, Milwaukee.
- Limp, W.F., K.L. Kvamme, E.G. Ernenwein, D.L. Harmon, M.L. Hargrave, T.L. Sever, and L.E. Somers (2002). *Multi-Dimensional Remote Sensing: A SERDP Project Fusing Ground, Air, and Satellite Data from Archaeological Sites*. Poster presented at the Partners in Environmental Technology Technical Symposium & Workshop, Department of Defense.
- Ernenwein, Eileen G., and Kenneth L. Kvamme (2002). *Multi-deimensional Remote Sensing at Army City, Kansas: A SERDP Project Fusing Ground, Air, and Satellite Data.* Paper presented at the 60th Annual Plains Anthropological Conference, Oklahoma City.

Published Book Chapters

- K.L. Kvamme, J.K. Johnson and B.S. Haley (2006). Integrating and Interpretation of Multiple Instrument Applications, In *Geophysical and Airborne Remote Sensing Applications in Archaeology: A Guide for Cultural Resource Managers*, J. Johnson, ed., University of Alabama Press, Tuscaloosa.
- Kvamme, Kenneth L. (2006). Archaeological Modeling with GIS at Scales Large and Small. In *Reading Historical Spatial Information from Around the World: Studies of Culture and Civilization Based on Geographic Information Systems Data.* Uno Takao, editor, International Research Center for Japanese Studies, Kyoto.
- Kvamme, Kenneth L. (2006). Integrating Multiple High Resolution Geophysical Data Sets. In *Remote Sensing in Archaeology,* J. Wiseman and F. El-Baz, editors, Plenum Publishers, New York, *volume in preparation.*
- Kvamme, Kenneth L. (2006). Remote Sensing: Archaeological Reasoning Through Physical Principles and Pattern Recognition, In *Archaeological Concepts for the Study of the Cultural Past*, A.P. Sullivan III, ed., University of Utah Press, Salt Lake City, *volume in preparation*.
- Tullis, Jason, Jack Cothren, Kenneth L. Kvamme (2006). In Situ Methods. In *Manual of Remote Sensing,* CRC Press, New York, *volume in preparation.*