ESTCP
Cost and Performance Report

(MM-0504)

Practical Discrimination Strategies for Application to Live Sites

November 2009

ENVIRONMENTAL SECURITY
TECHNOLOGY CERTIFICATION PROGRAM

U.S. Department of Defense

COST & PERFORMANCE REPORT
Project: MM-0504

TABLE OF CONTENTS

Page

1.0 EXECUTIVE SUMMARY 1

2.0 INTRODUCTION 3
- 2.1 BACKGROUND 3
- 2.2 OBJECTIVES OF THE DEMONSTRATION 3
- 2.3 REGULATORY DRIVERS 4

3.0 TECHNOLOGY 5
- 3.1 TECHNOLOGY DESCRIPTION 5
  - 3.1.1 Feature Extraction 6
  - 3.1.2 Discrimination Using Rule-Based or Statistical Classifiers 8
  - 3.1.3 UXOLab Software 11
  - 3.1.4 Previous Tests of the Technology 11
- 3.2 ADVANTAGES AND LIMITATIONS OF THE TECHNOLOGY 14

4.0 PERFORMANCE OBJECTIVES 15

5.0 SITE DESCRIPTION 17
- 5.1 FORMER LOWRY BOMBING AND GUNNERY RANGE 17
  - 5.1.1 Site Location and History 17
  - 5.1.2 Munitions Contamination 17
- 5.2 FORMER CAMP SIBERT 19
  - 5.2.1 Site Location and History 19
  - 5.2.2 Munitions Contamination 20
- 5.3 FORT MCCLELLAN 20
  - 5.3.1 Site Location and History 20
  - 5.3.2 Site Geology 22
  - 5.3.3 Munitions Contamination 22

6.0 TEST DESIGN 23
- 6.1 FORMER LOWRY BOMBING AND GUNNERY RANGE 23
  - 6.1.1 Conceptual Experimental Design 23
  - 6.1.2 Site Preparation 24
  - 6.1.3 System Specification 24
  - 6.1.4 Data Collection 27
  - 6.1.5 Validation 28
- 6.2 FORMER CAMP SIBERT 29
  - 6.2.1 Conceptual Experimental Design 29
  - 6.2.2 Site Preparation 30
  - 6.2.3 System Specification and Data Collection 31

TABLE OF CONTENTS (continued)

Page

6.3 FORT MCCLELLAN 31
6.3.1 Conceptual Experimental Design 31
6.3.2 Site Preparation 31
6.3.3 System Specification 32
6.3.4 Data Collection 33
6.3.5 Validation 36

7.0 DATA ANALYSIS AND PRODUCTS 37
7.1 PREPROCESSING 37
7.2 TARGET SELECTION FOR DETECTION 37
7.3 PARAMETER ESTIMATES 38
7.4 CLASSIFIER AND TRAINING 39
7.4.1 Classification Strategy at FLBGR 39
7.4.2 Classification Strategy at Camp Sibert 41
7.4.3 Classification Strategy at Fort McClellan 44
7.5 DATA PRODUCTS 45

8.0 PERFORMANCE ASSESSMENT 47
8.1 DISCRIMINATION PERFORMANCE AT FLBGR 47
8.2 PERFORMANCE METRICS AT CAMP SIBERT 50
8.2.1 Discrimination Performance of Full-Coverage Data Sets 53
8.2.2 Discrimination Performance of Cued-Interrogation Data Sets 54
8.2.3 Processing Time Required 56
8.3 PERFORMANCE METRICS AT FORT MCCLELLAN 57
8.3.1 Discrimination Performance at Fort McClellan 58
8.3.2 Number of Anomalies Covered Per Day 60

9.0 COST ASSESSMENT 61
9.1 COST SUMMARY 61
9.1.1 FLBGR Cost Summary 62
9.1.2 Camp Sibert Data Collection Summary 68
9.1.3 Fort McClellan Data Collection Summary 68
9.1.4 Projected Costs for Future EM63 Cued-Interrogation Deployments 70
9.2 CAMP SIBERT DISCRIMINATION 71

10.0 IMPLEMENTATION ISSUES 73

11.0 REFERENCES 75

APPENDIX A POINTS OF CONTACT A-1

LIST OF FIGURES

Page

Figure 1. Schematic of discrimination process. 6
Figure 2. Schematic of feature extraction process for EMI and magnetometer data. 8
Figure 3. A framework for statistical pattern recognition. 9
Figure 4. Nonparametric density estimate using Gaussian kernels. 10
Figure 5. Support vector machine formulation for constructing a decision boundary. 11
Figure 6. Locations of the RR and 20 mm RF sites at FLBGR. 18
Figure 7. Camp Sibert site map with initial magnetometer survey locations shown. 20
Figure 8. Fort McClellan site map. 21
Figure 9. Map of RR with areas surveyed for this demonstration outlined in red. 25
Figure 10. Map of the 20 mm RF, with two grids surveyed for this demonstration outlined in red. 26
Figure 11. Sky's EM 61-MK2 towed array, which is constructed of composite materials and houses EM sensors, RTS laser positioning (or GPS) sensors, and the Crossbow IMU. 27
Figure 12. Equipment used at FLBGR including the modified EM63 cart (left) and the Leica RTS TPS1206 laser positioning system (right). 27
Figure 13. Standard EM63 cart collecting discrimination mode data at the Ashland test site. 32
Figure 14. Surveying difficulties encountered on the site including (a) steep slopes, (b) rough surfaces, and (c) cramped survey areas due to closely spaced trees. 36
Figure 15. Scatterplot of the moment versus remenance for the (a) training data and (b) test data. 41
Figure 16. Feature vector plots for the contractor EM61 (a and b), MTADS EM61 (c and d), and MTADS EM61 cooperative (e and f) datasets. 42
Figure 17. Feature vector plots for the contractor EM63 (a and b) and EM63 cooperative (c and d) datasets. 43
Figure 18. Quadratic discriminant analysis classification of training and test data. 45
Figure 19. Ranked dig list. 46
Figure 20. Comparison of ROC curves for the magnetics (moment), contractor EM61, MTADS EM61, and MTADS EM61 cooperative data sets: (a) including can't-analyze category and (b) excluding can't-analyze anomalies. 54
Figure 21. ROC curves for the EM63 (a and b), EM63 cooperative (c and d), and MTADS EM61 cooperative (e and f) on each of the cued-interrogation anomalies. 55
Figure 22. ROC curves corresponding to the quadratic discriminant analysis classifier of Figure 18. 59
Figure 23. Comparison of time decay parameter feature spaces obtained from (a) inversion of a physics based model and (b) from decay curve analysis. 59
Figure 24. Number of cued interrogation anomalies surveyed each day. 60

LIST OF TABLES

Page

Table 1. Inversion/classification tests. 12
Table 2. Performance objectives for the FLBGR demonstration....15
Table 3. Performance objectives for Camp Sibert demonstration study....16
Table 4. Performance objectives for the Fort McClellan demonstration. 16
Table 5. Key project activities 27
Table 6. Number of anomalies identified as MK-23, 2.25-inch rocket, 37 mm projectile, 20 mm projectile, small arms, shrapnel, and junk in each of the 10 grids used for this demonstration. 29
Table 7. The number of items surveyed in each grid. 34
Table 8. Different phases of classification on the 20 mm RF and the RR and the corresponding training data and feature vectors selected for classification for the EM61 and EM63. 40
Table 9. Performance criteria and the metrics used for evaluation at FLBGR. 47
Table 10. Performance metrics at FLBGR....48
Table 11. EM61 discrimination performance results at the 20 mm RF and RR. 49
Table 12. EM63 discrimination performance results at both sites. 50
Table 13. Expected performance and performance confirmation methods. 51
Table 14. Summary of the discrimination performance of the nine different sensor combinations or discrimination methods. 52
Table 15. Time spent processing each of the different sensor combinations. 57
Table 16. Expected performance and performance confirmation methods. 58
Table 17. Cost summary for all project activities and demonstrations. 61
Table 18. Cost model for data collection demonstrations....62
Table 19. FLBGR demonstration survey costs. 64
Table 20. Assumptions used to compare the different survey methods....66
Table 21. Comparison of the costs for the different modes of survey using the assumptions in Table 17 and in the bullet points immediately before that table. 66
Table 22. Comparison of the cost of survey for the different methods with percentages of holes to dig and different numbers of anomalies per acre. 67
Table 23. Comparison of the cost of survey for the different methods with different percentages of holes to dig and different amounts of time required for interpretation of each anomaly....67
Table 24. Fully burdened costs for the Camp Sibert demonstration and the premobilization tests conducted in Ashland. 68
Table 25. Fort McClellan demonstration cost summary. 69
Table 26. Per anomaly cost breakdown. 70
Table 27. Future deployment cost estimates. 71
Table 28. Camp Sibert discrimination cost summary. 72

ACRONYMS AND ABBREVIATIONS

Am^2	ampere-meter squared
ASR	Archives Search Report
CDTF	Chemical Decontamination Training Facility
COBRA	Chemical, Ordnance, Biological and Radiological
COTS	commercial off-the-shelf
Cs	cesium
DoD	Department of Defense
EM	electromagnetic
EMI	electromagnetic induction
EOD	explosive ordnance disposal
ERDC	Engineer Research and Development Center
ESTCP	Environmental Security Technology Certification Program
FAR	false alarm rate
FLBGR	Former Lowry Bombing and Gunnery Range
FOM	Figure of Merit
FP	false positive
FUDS	Formerly Used Defense Site
GPO	geophysical prove-out
GPS	Global Positioning System
HE	high explosive
IDA	Institute for Defense Analyses
IMU	inertial measurement unit
IPR	in-progress review
m	meter
MEC	munitions and explosives of concern
mm	millimeter
MOUT	Military Operations in Urbanized Terrain
MRA	Munitions Response Area
MRS	Munitions Response Sites
MSEMS	Maintenance Standardization and Evaluation Management System
MTADS	multisensor towed array detection system
mV	millivolts
NAEVA	NAEVA Geophysics Inc.
nT	nanoTesla
OP	operating point

ACRONYMS AND ABBREVIATIONS (continued)

Pd	probability of detection
Pdisc	probability of discrimination
PDF	Portable Document Format
Pfa	probability of false alarm
PNN	Probabilistic Neural Network
QA	quality assurance
QC	quality control
RF	Range-Fan
ROC	receiver operating characteristic
RR	Rocket Range
RTC	Replacement Training Center
RTK-GPS	Real-Time-Kinematic Global Position System
RTS	robotic total station
SE1	South-East-1
SE2	South-East-2
SERDP	Strategic Environmental Research and Development Program
Sky	Sky Research, Inc.
SNR	signal-to-noise ratio
STOLS	Surface-Towed Ordnance Locator System
SVM	support vector machine
SW	South-West
TEM	time-domain electromagnetic
TOI	target of interest
TPS	Total Stations (Leica)
UBC	University of British Columbia
UBC-GIF	University of British Columbia-Geophysical Inversion Facility
USACE	U.S. Army Corps of Engineers
UTC	Unit Training Center
UXO	unexploded ordnance
WP	white phosphorus
YPG	Yuma Proving Ground

ACKNOWLEDGEMENTS

Sky Research, Inc. (Sky) performed this work for the Environmental Security Technology Certification Program (ESTCP) under Project MM-200504.

The work described in this demonstration report was performed jointly by Sky and the University of British Columbia-Geophysical Inversion Facility (UBC-GIF). All work was conducted at either the Sky Vancouver office, or at the UBC-GIF office which is also in Vancouver. Dr. Stephen Billings served as Principal Investigator.

Funding for this project was provided by the ESTCP Program Office. This project offered the opportunity to demonstrate this technology and evaluate its potential to support the Department of Defense's (DoD) efforts to discriminate hazardous ordnance from non-hazardous shrapnel, ordnance related scrap, and cultural debris.

We wish to express our sincere appreciation to Dr. Jeffrey Marqusee, Dr. Anne Andrews, Dr. Herbert Nelson, and Ms. Katherine Kaye of the ESTCP Program Office for providing support and funding for this project, and for facilitating access to the technical information required to complete this demonstration. We also would like to acknowledge the support and direction provided by Mr. Don Yule of the U.S. Army Corps of Engineers-Engineer Research and Development Center (USACE-ERDC), the Contracting Officer's Representative for this project.

We also would like to acknowledge the long-term support and funding provided by the USACE-ERDC and Dr. John Cullinane in particular. The work reported here is a demonstration of techniques largely funded through the USACE-ERDC.

The work presented herein was accomplished between September 2005 and November 2009.

Technical material contained in this report has been approved for public release.

This page left blank intentionally.

1.0 EXECUTIVE SUMMARY

This project addressed one of the Department of Defense's (DoD) most pressing environmental problems—the efficient and reliable identification of unexploded ordnance (UXO) without the need to excavate large numbers of non-ordnance items. Electromagnetic (EM) sensors have been shown to be a very promising technology for detecting UXO, but they also tend to detect many other nonhazardous metallic items. Current cleanup practice is to excavate all anomalies with peak amplitude above a predefined threshold. Such techniques are inefficient and costly, with at times over 100 nonhazardous items excavated for each UXO. Much research over the past few years has been focused on the discrimination problem whereby features from physics-based model-fits to anomalies are used to determine the likelihood that the buried item is a UXO. Statistical and rule-based classification techniques, when calibrated with good training data, have been shown at numerous test-trial sites to be very effective at discrimination. However, guidelines and standard operating procedures for their application to live sites have yet to be established. The principal objectives of the work conducted here were to develop a practical strategy for discrimination of UXO that can be reliably applied to real sites along with the protocols and tools to test performance. Three different demonstrations were conducted under this project. The first demonstration of the methodology was conducted at the Former Lowry Bombing and Gunnery Range (FLBGR) in Colorado during the 2006 field season. The focus of the FLBGR demonstration was on verification of the single inversion process used to extract physics-based parameters from magnetic and electromagnetic induction (EMI) anomalies, and on the statistical classification algorithms used to make discrimination decisions from those parameters. Two sites were visited at FLBGR. The Rocket Range (RR) survey objective was to discriminate a mixed range of projectiles with a minimum diameter of 37 mm from shrapnel, junk, 20 mm projectiles, and small arms. The 20 mm Range-Fan (RF) survey presented a small-item discrimination scenario with a survey objective of discriminating 37 mm projectiles from ubiquitous 20 mm projectiles and 50-caliber bullets.

At the FLBGR site, two phases of digging and training were conducted at the 20 mm RF and three phases at the RR. At the RR, 29 MK-23 practice bombs were recovered, with only one other UXO item encountered (a 2.5 inch rocket warhead). At the 20 mm RF, 38 37 mm projectiles (most of them emplaced) were recovered, as were a large number of 20 mm projectiles and 50-caliber bullets. For both sites, and for both instruments, the support vector machine (SVM) classifier outperformed a ranking based on amplitude alone. In each case, the last detected UXO was ranked quite high by the SVM classifier, and digging to that point would have resulted in a 60-90% reduction in the number of false alarms. This operating point is of course unknown prior to digging. We found that using a stop-digging criteria of f=0 (midway between UXO and clutter class support planes) was too aggressive and more excavations were typically required for full recovery of detected UXO. Both the amplitude and SVM methods performed quite poorly on two deep (40 cm) emplaced 37 mm projectiles at the 20 mm RF, exposing a potential weakness of the goodness-of-fit metric. Retrospective analysis revealed that thresholding on the size of the polarization tensor alone would have yielded good discrimination performance.

The second demonstration was conducted as part of the Environmental Security Technology Certification Program (ESTCP) discrimination pilot study at Camp Sibert, AL, during 2007. The

objective was to find potentially hazardous 4.2-inch mortars. The demonstration provided another test of the methodology as well as that of the cooperative inversion process. Both cued interrogation and full coverage data collected by different demonstrators were analyzed, allowing the effect of data quality on discrimination decisions to be assessed. For the Camp Sibert discrimination study, the project team created eight different dig-sheets from six different sensor combinations: (1) multisensor towed array detection system (MTADS) magnetics; (2) EM61 cart (classification and size based); (3) MTADS EM61 (classification and size based); (4) MTADS EM61 and magnetics; (5) EM63; and (6) EM63 and magnetics.

The results for all sensor combinations were excellent, with just one false negative for the EM63 when inverted without cooperative constraints. When inverted cooperatively, the EM63 cued interrogation was the most effective discriminator. All 33 UXO were recovered with 25 false alarms (16 of these were in the "can't-analyze" category). Not counting the can't-analyze category, the first 33 recommended excavations were all UXO. The MTADS and MTADS cooperatively inverted were also very effective at discrimination, with all UXO recovered very early in the dig list (e.g., for the MTADS cooperative there were just two false positives (FP) by the time all 117 can't-analyze UXO were recovered). The MTADS data set suffered from a high number of false alarms due to anomalies with a geological origin (caused by the cart bouncing up and down). In addition, the operating point was very conservative and many non-UXO were excavated after recovery of the last UXO in the dig list. The results from the EM61 cart were also very good, although 24 FPs were required to excavate all 105 UXO (that weren't in the can't-analyze category). The lower data quality of the EM61 cart resulted in a larger number of can't-analyze anomalies over metallic sources than the MTADS.

The objectives of the third demonstration were to evaluate the discrimination potential of the Geonics EM63 at Fort McClellan, AL, when deployed in a cued interrogation mode. Pasion-Oldenburg polarization tensor models were fit to each of the EM63 cued anomalies. Ground truth information from 60 of the 401 live-site anomalies, along with 18 items in the geophysical prove-out and 21 items measured in a test pit were used to train a statistical classifier. Features related to shape, encapsulated in the relative values of the primary, secondary, and tertiary polarizations, were unstable and could not be used for reliable discrimination. A feature space comprising the size and the relative decay rate of the primary polarization was used for discrimination of the medium caliber projectiles (75 mm and 3.8-inch shrapnel).

All demonstration metrics related to discrimination of these medium caliber projectiles were met. At the operating point, all but five of 119 targets of interest were recommended for excavation, with 34 false alarms. If the operating point was relaxed slightly, then all medium caliber projectiles would have been recovered with 51 false alarms. Retrospective analysis revealed that excellent discrimination performance could have been obtained by using a feature space comprising an early and late time feature extracted from the object's primary polarization. Furthermore, we found that these feature vectors could be approximated without fitting polarization tensor models to the data and by using just seven measurement locations around the template center. These approximate early and late time decay features were extracted from the sounding with the slowest decay (defined as the ratio of the 20th to 1st time channels).

2.0 INTRODUCTION

2.1 BACKGROUND

The fiscal year 2006 (FY06) Defense Appropriation contains funding for the *Development of Advanced, Sophisticated, Discrimination Technologies for UXO Cleanup* in ESTCP. In 2003, the Defense Science Board observed: "The … problem is that instruments that can detect the buried UXO also detect numerous scrap metal objects and other artifacts, which leads to an enormous amount of expensive digging. Typically 100 holes may be dug before a real UXO is unearthed! The Task Force assessment is that much of this wasteful digging can be eliminated by the use of more advanced technology instruments that exploit modern digital processing and advanced multi-mode sensors to achieve an improved level of discrimination of scrap from UXO." Significant progress has been made in discrimination technology development. To date, testing of these approaches has been primarily limited to test sites with only limited application at live sites. Acceptance of discrimination technologies requires demonstration of system capabilities at real UXO sites under real world conditions. Any attempt to declare detected anomalies to be harmless and requiring no further investigation will require demonstration to regulators of not only individual technologies but an entire decision making process.

2.2 OBJECTIVES OF THE DEMONSTRATION

Three different demonstrations were conducted under this project. The first demonstration of the methodology was conducted at the Former Lowry Bombing and Gunnery Range (FLBGR) in Colorado during the 2006 field season. The focus of the FLBGR demonstration was on the verification of the single inversion process used to extract physics-based parameters from magnetic and EMI anomalies, as well as the statistical classification algorithms used to make discrimination decisions from those parameters. The second demonstration was conducted as part of the ESTCP discrimination pilot study at Camp Sibert, AL, during 2007. The objective was to find potentially hazardous 4.2-inch mortars. The demonstration provided another test of the methodology as well as that of the cooperative inversion process. Both cued interrogation and full coverage data collected by different demonstrators were analyzed, allowing the effect of data quality on discrimination decisions to be assessed. For the Camp Sibert discrimination study, the project team created eight different dig sheets from six different sensor combinations: (1) MTADS magnetics; (2) EM61 cart (classification and size based); (3) MTADS EM61 (classification and size based); (4) MTADS EM61 and magnetics; (5) EM63; and (6) EM63 and magnetics. Effective discrimination was demonstrated for all sensor combinations, with just one false negative for the EM63 when inverted without magnetometer location constraints. The cued interrogation EM63 data when cooperatively inverted with the magnetics data was the most effective discriminator.

The objectives of the third demonstration were to evaluate the discrimination potential of the Geonics EM63 at Fort McClellan, AL, when deployed in a cued interrogation mode. Pasion-Oldenburg polarization tensor models were fit to each of the EM63 cued anomalies. Feature vectors extracted from those dipole fits were used to guide a statistical classification algorithm that ranked the items in order of UXO likelihood.

2.3 REGULATORY DRIVERS

The Defense Science Board Task Force on UXO noted in its 2003 report that 75% of the total cost of a current clearance is spent on digging scrap. A reduction from 100 to 10 in the number of scrap items dug per UXO item could reduce total clearance costs by as much as two-thirds. Thus, discrimination efforts focus on technologies that can reliably differentiate UXO from items that can be safely left undisturbed.

Discrimination becomes a realistic option only when the cost of identifying items that may be left in the ground is less than the cost of digging them. Because discrimination requires detection as a precursor step, the investment in additional data collection and analysis must result in enough fewer items dug to pay back the investment. Even with perfect detection performance and high signal-to-noise ratio (SNR) values, successfully sorting the detections into UXO and nonhazardous items is a difficult problem but, because of its potential payoff, one that is the focus of significant current research. The demonstrations conducted under this project represent an effort to transition a promising discrimination technology into widespread use at UXO-contaminated sites across the country.

3.0 TECHNOLOGY

3.1 TECHNOLOGY DESCRIPTION

Magnetic and EM methods represent the main sensor types used for detection of UXO. Over the past 10 years, significant research effort has been focused on developing methods to discriminate between hazardous UXO and nonhazardous scrap metal, shrapnel and geology (e.g., Hart et al., 2001; Collins et al., 2001; Pasion & Oldenburg, 2001; Zhang et al., 2003a, 2003b; Billings, 2004). The most promising discrimination methods typically proceed by first recovering a set of parameters that specify a physics-based model of the object being interrogated. For example, in time-domain electromagnetic (TEM) data, the parameters comprise the object location and the polarization tensor (typically two or three collocated orthogonal dipoles along with their orientation and some parameterization of the time decay curve). For magnetics, the physics based model is generally a static magnetic dipole. Once the parameters are recovered by inversion, a subset of the parameters is used as feature vectors to guide either a statistical or rule-based classifier.

Magnetic and EM phenomenologies have different strengths and weaknesses. Magnetic data are simpler to collect, are mostly immune to sensor orientation, and are better able to detect deeper targets. EM data are sensitive to nonferrous metals, are better at detecting smaller items and are able to be used in areas with magnetic geology. Therefore, there are significant advantages in collecting both types of data including increased detection, stabilization of the EM inversions by cooperative inversion of the magnetics (Pasion et al., 2003), and extra dimensionality in the feature space that may improve classification performance (e.g., Zhang et al., 2003a). However, these advantages need to be weighed against the extra costs of collecting both data types.

There are three key elements that impact the success of the UXO discrimination process described in the previous paragraphs (Figure 1):

(1) *Collection of data and creation of a map of the geophysical sensor data.* This includes all actions required to form an estimate of the geophysical quantity in question (magnetic field in nanoTesla [nT], amplitude of EMI response at a given time channel, etc.) at each of the visited locations. The estimated quantity is dependent on the following:

- Hardware, including the sensor type, deployment platform, position and orientation system, and the data acquisition system used to record and time-stamp the different sensors
- Survey parameters such as line spacing, sampling rate, calibration procedures etc.
- Data processing such as merging of position/orientation information with sensor data, noise, and background filtering applied
- The background environment including geology, vegetation, topography, cultural features, etc.
- Depth and distribution of ordnance and clutter.

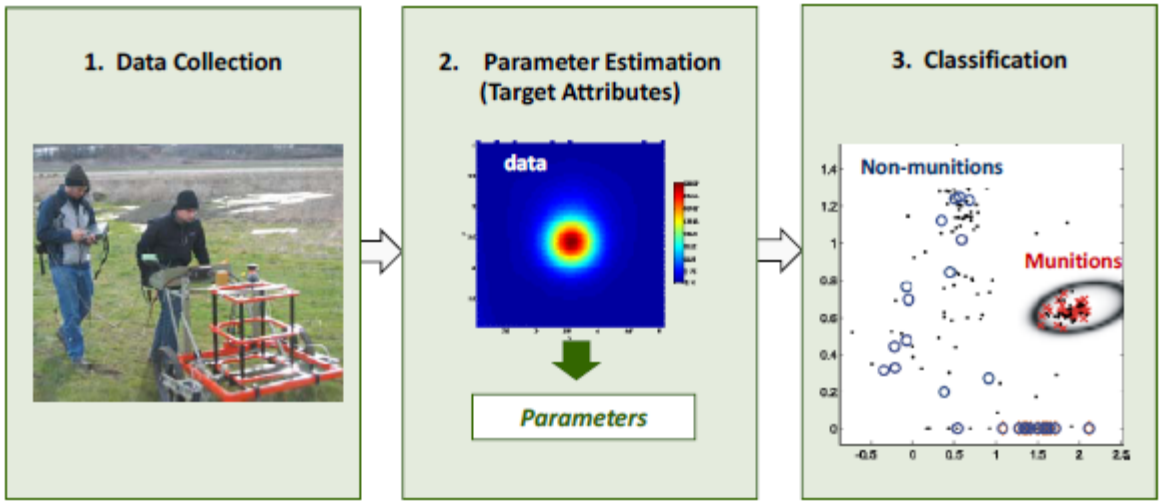

Figure 1. Schematic of discrimination process.

(2) *Anomaly selection and feature extraction.* This includes the detection of anomalous regions and the subsequent extraction of a dipole (magnetics) or polarization tensor (TEM) model for each anomaly. Where magnetic and EMI data have both been collected, the magnetic data can be used as constraints for the EMI model via a cooperative inversion process.

(3) *Classification of anomalies.* The final objective of the demonstration is the production of a dig sheet with a ranked list of anomalies. This will be achieved via statistical classification that will require training data to determine the attributes of the UXO and non-UXO classes.

The focus of demonstrations conducted under this project was on the validation of the methodologies for (2) and (3) above that have been developed in UXOLab jointly by Sky Research, Inc. (Sky) and the University of British Columbia-Geophysical Inversion Facility (UBC-GIF). The success of the discrimination process will be critically dependent on the attributes of the data used for the feature extraction and subsequent classification (vis-a-vis, everything pertaining to the first element described above), in particular, the SNR, location accuracy, sampling density and information content of the data (the more time channels or vector components, the more information that will be available to constrain the fits). Thus, while our intent is to test the algorithms developed in UXOLab, this test cannot be conducted in isolation of the attributes of the geophysical sensor data.

3.1.1 Feature Extraction

In the EMI method, a time varying field illuminates a buried, conductive target. Currents induced in the target then produce a secondary field that is measured at the surface. EM data inversion involves using the secondary field generated by the target for recovery of the position, orientation, and parameters related to the target's material properties and shape. In the UXO community, the inverse problem is simplified by assuming that the secondary field can be accurately approximated as a dipole.

In general, TEM sensors use a step-off field to illuminate a buried target. The currents induced in the buried target decay with time, generating a decaying secondary field that is measured at the surface. The time-varying secondary magnetic field $\mathbf{B}(t)$ at a location $\mathbf{r}$ from the dipole $\mathbf{m}(t)$ is:

$$\mathbf{B}(t) = \frac{\mu_o}{4\pi r^3}\mathbf{m}(t)\cdot\left(3\hat{\mathbf{r}}\hat{\mathbf{r}} - \overline{\overline{\mathbf{I}}}\right) \tag{1}$$

where $\hat{\mathbf{r}} = \mathbf{r}/|\mathbf{r}|$ is the unit-vector pointing from the dipole to the observation point, I is the 3 x 3 identity matrix, $\mu_o = 4\ \pi \times 10^{-7}$ H/m is the permittivity of free space and $r = |\mathrm{r}|$ is the distance between the center of the object and the observation point.

The dipole induced by the interaction of the primary field $\mathbf{B}_o$ and the buried target is given by:

$$\mathbf{m}(t) = \frac{1}{\mu_o}\overline{\overline{\mathbf{M}}}(t)\cdot\mathbf{B}_o \tag{2}$$

where $\mathbf{M}(t)$ is the target's polarization tensor. The polarization tensor governs the decay characteristics of the buried target and is a function of the shape, size, and material properties of the target. The polarization tensor is written as:

$$\overline{\overline{\mathbf{M}}}(t) = \begin{bmatrix} L_1(t) & 0 & 0 \\ 0 & L_2(t) & 0 \\ 0 & 0 & L_3(t) \end{bmatrix} \tag{3}$$

where we use the convention that $L_1(t_1) \geq L_2(t_1) \geq L_3(t_1)$, so that polarization tensor parameters are organized from largest to smallest.

Given the TEM data measured over an anomaly, the objective of the feature extraction process is the accurate estimation of the polarization tensor parameters. This is achieved by finding the polarization tensor model (including location, depth and orientation) that best matches the observed data. If both magnetometer and TEM data are available, the feature extraction can be achieved using "cooperative inversion." In that process, the magnetometer data are analyzed first and the location and depth of the recovered magnetic dipole model can be used to constrain the location and depth of the TEM polarization tensor model. This helps to stabilize the TEM feature extraction by minimizing the ambiguity in the location and depth of the polarization tensor model.

A schematic of the feature extraction process is shown in Figure 2.

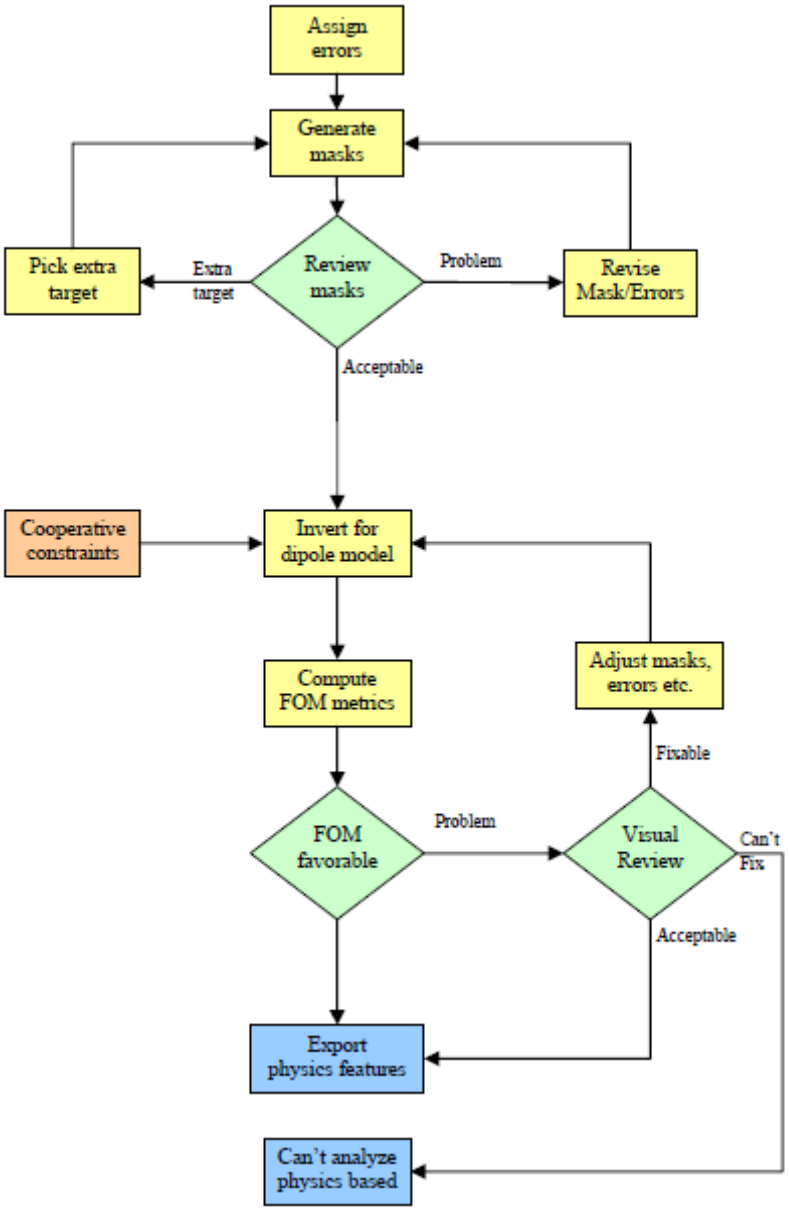

Figure 2. Schematic of feature extraction process for EMI and magnetometer data.

3.1.2 Discrimination Using Rule-Based or Statistical Classifiers

At this stage in the process, we have feature vectors for each anomaly and now need to decide which items should be excavated as potential UXO. Rule-based classifiers use relationships derived from the underlying physics to partition the feature space. Examples include the ratio of TEM decay parameters (Pasion and Oldenburg, 2001) and magnetic remanence (Billings, 2004). For this demonstration, we focused on statistical classification techniques which have proven to be very effective at discrimination at various test sites (e.g., Zhang et al., 2003b).

Statistical classifiers have been applied to a wide variety of pattern recognition problems, including optical character recognition, bioinformatics and UXO discrimination. Within this field there is an important dichotomy between "supervised" and "unsupervised" classification. Supervised classification makes classification decisions for a test set consisting of unlabeled feature vectors. The classifier performance is optimized using a training data set for which labels are known. In unsupervised classification there is only a test set; labels are unknown for all feature vectors. Most applications of statistical classification algorithms to UXO discrimination have used supervised classification; the training data set is generated as targets are excavated. More recently, unsupervised methods have been used to generate a training data set that an informative sample of the test data (Carin et al., 2004). In addition, "semi-supervised" classifiers, which exploit both labeled data and the topology of unlabeled data, have been applied to UXO discrimination in one study (Carin et al., 2004).

Figure 3 summarizes the supervised classification process within the statistical framework. Given test and training data sets, we extract features from the data, select a relevant subset of these features and optimize the classifier using the available training data. Because the predicted performance of the classifier is dependent upon the feature space, the learning stage can involve further experimentation with feature extraction and selection before adequate performance is achieved.

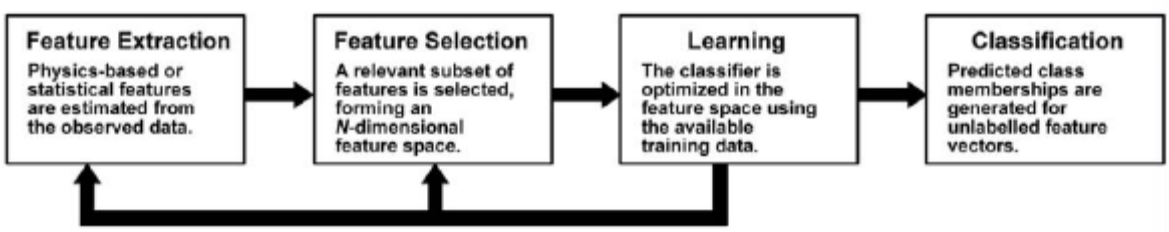

Figure 3. A framework for statistical pattern recognition.

There are two (sometimes equivalent) approaches to partitioning the feature space. The generative approach models the underlying probability distributions that are assumed to have produced the observed feature data. The starting point for any generative classifier is Bayes rule:

$$\mathrm{P}(\omega_i \,|\mathbf{x}) \,\alpha\, \mathrm{P}(\mathbf{x}|\, \omega_i)\mathrm{P}(\omega_i). \tag{4}$$

The likelihood function $\mathrm{P}(\mathbf{x}|\, \omega_i)$ computes the probability of observing the feature vector **x** given the class ω_i. The prior probability $\mathrm{P}(\omega_i)$ quantifies our expectation of how likely we are to observe class ω_i. Bayes rule provides a mechanism for classifying test feature vectors: assign **x** to the class with the largest *a posteriori* probability. Contours along which the posterior probabilities are equal define decision boundaries in the feature space.

An example of a generative classifier is discriminant analysis, which assumes a Gaussian form for the likelihood function. Training this classifier involves estimating the means and co-variances of each class. If equal co-variances are assumed for all classes, the decision boundary

is linear. While these assumptions may seem overly restrictive, in practice linear discriminant analysis performs quite well in comparison with more exotic methods and is often used as a baseline classifier when assessing performance.

Other generative classifiers assume a nonparametric form for the likelihood function. For example, the Probabilistic Neural Network (PNN) models the likelihood for each class as a superposition of kernel functions. The kernels are centered at the training data for each class. In this case, the complexity of the likelihood function (and hence the decision boundary) is governed by the width of the kernels (Figure 4).

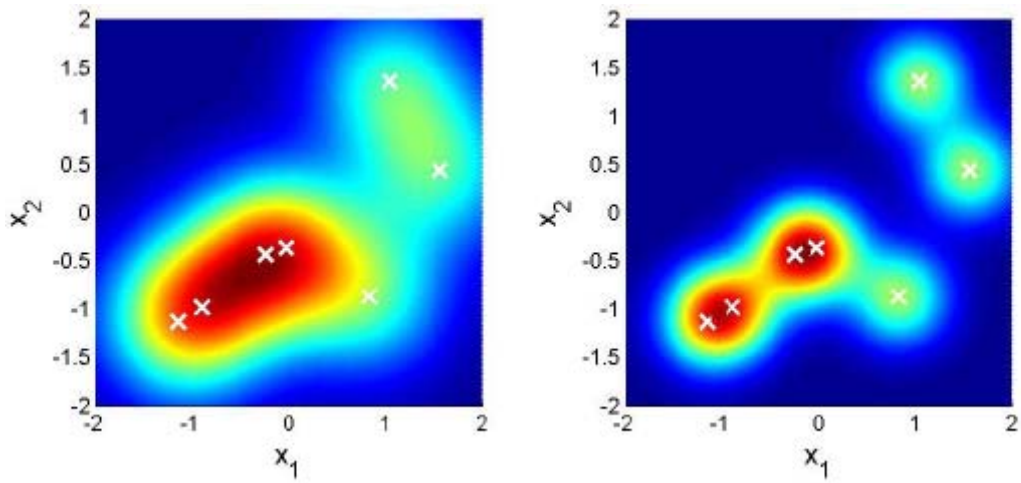

Figure 4. Nonparametric density estimate using Gaussian kernels.
Kernel centers are shown as crosses. A large kernel width produces a smooth distribution (left) compared to a small kernel width (right).

The discriminative approach is not concerned with underlying distributions but rather seeks to identify decision boundaries that provide an optimal separation of classes. For example, a support vector machine (SVM) constructs a decision boundary by maximizing the margin between classes. The margin is defined as the perpendicular distance between support planes by which the classes are bound, as shown in Figure 5. The decision boundary then bisects the support planes. This formulation leads to a constrained optimization problem: to maximize the margin between classes subject to the constraint that the training data are classified correctly. An advantage of the SVM method over other discriminative classifiers (e.g. neural networks) is that there is a unique solution to the optimization problem.

With all classification algorithms, a balance must be struck between obtaining good performance on the training data and generalizing to a test data set. An algorithm that classifies all training data correctly may produce an overly complex decision boundary that may not perform well on the test data. In the literature this is referred to as “bias-variance trade-off” and is addressed by constraining the complexity of the decision boundary (regularization). In cases such as linear discriminant analysis, the regularization is implicit in specification of the likelihood function. Alternatively, the complexity of the fit can be explicitly governed by regularization parameters (e.g. the width of kernels in a PNN or Lagrange multipliers in an SVM). These parameters are typically estimated from the training data using cross-validation, which sets aside a portion of the

training data to assess classifier performance for a given regularization. We obtained our training data from the geophysical prove-out (GPO) and from the release of data over a minimum of 50 anomalies on the live site. Figure 5 illustrates the SVM formulation for constructing a decision boundary.

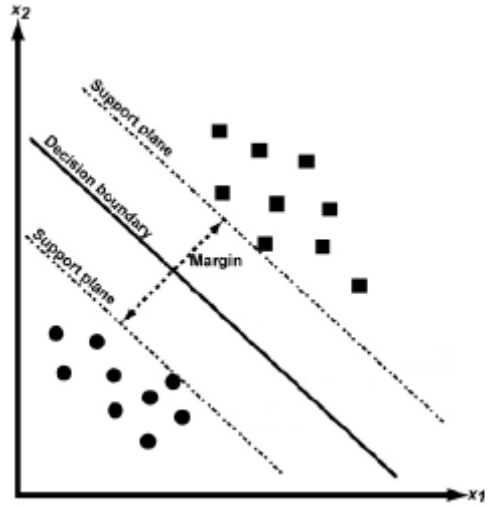

Figure 5. Support vector machine formulation for constructing a decision boundary. The decision boundary bisects support planes bounding the classes.

3.1.3 UXOLab Software

The methodologies for data processing, feature extraction, and statistical classification described above have been implemented within the UXOLab software environment. This is a Matlab based software package developed over a 7-year period at the UBC-GIF, principally through funding by the USACE Engineer Research and Development Center (ERDC) project (DAAD19-00-1-0120). Over the past 4 years, Sky and UBC-GIF have expanded the capabilities of the software considerably. These improvements have been largely sponsored by this ESTCP project, as well as several Strategic Environmental Research and Development Program (SERDP)-funded research projects.

3.1.4 Previous Tests of the Technology

Table 1 provides a summary of different sites where the technology has been tested, including the three sites visited as part of this project.

Table 1. Inversion/classification tests.

Inversion/Classification Test	Description	Results
Demonstration Site: Yuma Proving Ground		
Proof-of-concept of cooperative inversion, Yuma Proving Ground (YPG)	Test of cooperative inversion on EM63 and magnetometer data collected in 2003. TEM inversions used two decaying orthogonal dipoles, constrained using magnetics data. Three different classifiers (linear and quadratic discriminant analysis, and PNN) were applied to the cooperative inversion results.	Classification of cooperatively inverted data is easier than inversion w/o magnetic constraints. Cleaner separation of classes is achieved for k parameters recovered from cooperative inversion; single and cooperative inversion results are similar for β parameters. This test demonstrated the UXOLab capability to perform both cooperative inversion and statistical classification.
Demonstration Site: Aberdeen Proving Ground		
Geocenters Surface-Towed Ordnance Locator System (STOLS) EM61 and magnetometer data	Discrimination ability of the system was marginal due to the following: limitations in positional accuracy (5-10 cm), which is inadequate for advanced discrimination), lack of sensor orientation data, and low SNR. No statistical classification algorithms were applied.	Results contributed to the decision to enhance Sky sensor systems by including the use of RTS for positioning and inertial measurement unit (IMU) for sensor orientation. Demonstrated the feasibility of cooperative inversion of large volumes of data with UXOLab.
Demonstration Site: FLBGR RR (8 acres surveyed) and 20mm RF (2 acres surveyed)		
Geonics EM61 and EM63 single inversion, positioned by a Leica Total Stations (TPS) 1206 robotic total station (RTS) with orientation provided by a Crossbow AHRS 400 IMU. The RR survey objective was to discriminate a mixed range of projectiles with minimum diameter of 37 mm from shrapnel, junk, 20 mm projectiles and small-arms. The 20 mm RF survey presented a small-item discrimination scenario with survey objective of discriminating 37 mm projectiles from ubiquitous 20 mm projectiles and 50-caliber bullets.	For the EM61, 3-dipole instantaneous amplitude models were fit to the available 4 time channels, while for the EM63, 3-dipole Pasion-Oldenburg models were recovered from the 26 time channel data. Parameters of the dipole model were used to guide a statistical classification. Canonical and visual analysis of feature vectors extracted from the test plot data indicated that discrimination could best proceed using a combination of a size and a goodness-of-fit-based feature vector. An SVM classifier was then implemented based on those feature vectors and using the available training data.	Two phases of digging and training were conducted at the 20 mm RF and three phases at the RR. At the RR, 29 MK-23 practice bombs were recovered, with only one other UXO encountered (a 2.5-inch rocket warhead). At the 20 mm RF, 38 37 mm projectiles (most of them emplaced) were recovered, as were a large number of 20 mm projectiles and 50-caliber bullets. For both sites and for both instruments, the SVM classifier outperformed a ranking based on amplitude alone. In each case, the last detected UXO was ranked quite high by the SVM classifier, and digging to that point would have resulted in a 60-90% reduction in the number of false alarms. This operating point is, of course, unknown prior to digging. We found that using a stop-digging criteria of f=0 (midway between UXO and clutter class support planes) was too aggressive and more excavations were typically required for full recovery of detected UXO. Both the amplitude and SVM methods performed quite poorly on two deep (40 cm) emplaced 37 mm projectiles at the 20 mm RF, exposing a potential weakness of the goodness-of-fit metric. Retrospective analysis revealed that thresholding on the size of the polarization tensor alone would have yielded good discrimination performance.

Table 1. Inversion/classification tests. (continued)

Inversion/Classification Test	Description	Results
Demonstration Site: Camp Sibert		
Geonics EM61 cart, MTADS EM61 array, MTADS mag array, and EM63 single and cooperative inversions. EM63 cued interrogations were positioned by a Leica TPS 1206 RTS with orientation information provided by a Crossbow AHRS 400 IMU. The objective of the surveys was the discrimination of a large target (4.2-inch mortars). The site was unusual in that the primary munitions item known to have been used was the 4.2-inch mortar, thus providing a site where the discrimination is a case of identifying a single large target amongst smaller pieces of mortar debris and clutter.	For the EM61, 3-dipole instantaneous amplitude models were fit to the available 3 time channels, while for the EM63, 3-dipole Pasion-Oldenburg models were recovered from the 26 time channel data. MTADS and EM63 data were also cooperatively inverted. Parameters of the dipole model were used to guide a statistical classification.	The results for all sensor combinations were excellent, with just one false negative for the EM63 when inverted without cooperative constraints. When inverted cooperatively, the EM63 cued interrogation was the most effective discriminator. All 33 UXO were recovered with 25 false alarms (16 of these were in the can't-analyze category). Not counting the can't-analyze category, the first 33 recommended excavations were all UXO. The MTADS and MTADS cooperatively inverted were also very effective at discrimination, with all UXO recovered very early in the dig list (e.g., for the MTADS cooperative, there were just 2 FPs by the time all 117 can't-analyze UXO were recovered). The MTADS data set suffered from a high number of false alarms due to anomalies with a geological origin (caused by the cart bouncing up and down). In addition, the operating point was very conservative and many non-UXO were excavated after recovery of the last UXO in the dig list. The results from the EM61 cart were also very good, although 24 FPs were required to excavate all 105 UXO (that weren't in the can't-analyze category). The lower data quality of the EM61 cart resulted in a larger number of can't-analyze anomalies over metallic sources than the MTADS.
Demonstration Site: Fort McClellan		
Geonics EM63 deployed in a cued interrogation model demonstrated. A wide range of potential items of interest of different calibers included grenades, 37 mm projectiles, 60 mm mortars, 75 mm shrapnel and 3.8-inch shrapnel rounds. The EM63 surveys were cued off production mode EM61 data. A template (constructed from a sturdy pool liner) was centered over each anomaly and data were then collected at 55 pre-marked station locations distributed about the center of the template.	Polarization tensor models were fit to each surveyed anomaly. Ground truth information from 60 of the 401 live site anomalies, along with 18 items in the geophysical prove-out and 21 items measured in a test pit were used to train a statistical classifier. Features related to shape, encapsulated in the relative values of the primary, secondary, and tertiary polarizations, were unstable and could not be used for reliable discrimination. A feature space comprising the size and the relative-decay rate of the primary polarization was used for discrimination of the medium caliber projectiles (75 mm and 3.8-inch shrapnel).	All demonstration metrics related to discrimination of these medium caliber projectiles were met. At the operating point, all but 5 of 119 targets of interest were recommended for excavation, with 34 false alarms. If the operating point was relaxed slightly, then all medium caliber projectiles would have been recovered with 51 false alarms. Retrospective analysis revealed that excellent discrimination performance could have been obtained by using a feature space comprising an early and late time feature extracted from the object's primary polarization. Furthermore, we found that these feature vectors could be approximated without fitting polarization tensor models to the data and by using just seven measurement locations around the template center. These approximate early features and a late time decay feature were extracted from the sounding with the slowest decay (defined as the ratio of the 20th to 1st time channels).

3.2 ADVANTAGES AND LIMITATIONS OF THE TECHNOLOGY

The main advantages of the technology are a potential reduction in the number of nonhazardous items that need to be excavated, thus reducing the costs of UXO remediation. There are two key aspects to the demonstrated technology (1) hardware and (2) software. On the hardware side, we concentrated on the demonstration of commercial off-the-shelf (COTS) sensors like the EM61, EM63, and cesium (Cs) vapor magnetometer. As each of these instruments measure only one component of a vector field, a measurement at a single location provides limited information. As a consequence, relatively dense two-dimensional measurements are required for accurate recovery of relevant target parameters. These measurements must be very precisely positioned and oriented for discrimination to be successful. The SERDP and ESTCP have sponsored the development of a new generation of EMI instruments that have much improved capabilities relative to the sensors demonstrated under this project. We are in the process of demonstrating our processing and interpretation approach using these new generation sensors at the ESTCP demonstration site at San Luis Obispo, CA.

On the software side, advantages of UXOLab and the algorithms within the package include:

- The software contains all the functionality required to process raw geophysical data, detect anomalous regions, and perform geophysical inversion and discrimination.
- UXOLab contains algorithms for inverting magnetic and TEM data sets both separately and cooperatively using a number of different polarization tensor formulations.
- Has an extensive set of algorithms for rule-based and statistical classification algorithms.
- UXOLab has been configured in a modular fashion, so that as new sensor technologies come online (e.g., new TEM systems with multicomponent receivers, etc.), the inversion functionality will be immediately available to those new sensor systems.

While UXOLab is available under license from the University of British Columbia (UBC), it is not suitable for general distribution to government contractors. Firstly, using the software successfully requires advanced knowledge of geophysical inversion and statistical classification. Secondly, while the software doesn't require the user to have a Matlab license, it was built entirely within the Matlab software environment to support the needs of UXO researchers. Thirdly, UBC is not set up to provide maintenance and support for the software.

4.0 PERFORMANCE OBJECTIVES

The performance objectives established for each of the demonstration project sites are listed in Table 2 through Table 4.

Table 2. Performance objectives for the FLBGR demonstration.

Type of Performance Objective	Primary Performance Criteria	Expected Performance (Metric)	Actual Performance (Objective Met?)			
			EM61 RR	EM61 20mm RF	EM63 RR	EM63 20mm RF
Qualitative	Terrain/vegetation restrictions	Operator acceptance for use at the site	Yes	Yes	Met	Met
	Ease of use (hardware)	Operator and site geophysicist acceptance	Yes	Yes	Partly	Partly
Quantitative	Probability of detection (Pd) of EM63 sensor	≥ Pd for EM61 towed array	NA	NA	Unknown	Unknown
	Probability of discrimination (Pdisc) with a 50% reduction in false alarms	≥ 0.9	Yes	Yes	Yes	Yes
	False alarm rate with PDisc = 1	> 25% reduction in false alarms	Yes	Yes	Yes	Yes
	Location accuracy of interpreted items	<0.2m	Yes	No	Yes	Yes
	Survey rate for magnetometer system	1 hectare/day	Yes	Yes	Yes	Yes
	Survey rate for EM63 system	1/3 hectare/day	NA	NA	Yes	Yes
	Percent site coverage	>95%	Yes	Yes	Yes	Yes
	Processing time (initial processing)	< 1 day per tile (1 acre)	Yes	Yes	Yes	Yes
	Processing time (interpretation)	< 5 min operator time per anomaly	Unknown	Unknown	Unknown	Unknown
	Accuracy of inversion parameters	Within class variance of cooperative < single inversion	NA	NA	NA	NA

Table 3. Performance objectives for Camp Sibert demonstration study.

Type of Performance Objective	Primary Performance Criteria	Expected Performance (Metric)	Actual Performance (Objective Met?)
Quantitative	Pdisc on recovered items at selected operating point	>0.95	Not applicable (NA) as all 4.2-inch mortars were seeded
	Pdisc on emplaced items at selected operating point	> 0.95	Yes for all technologies
	False alarm rate with PDisc (recovered) = 0.95	> 50% reduction in false alarms	Yes for all technologies
	False alarm rate with PDisc = 1	> 25% reduction in false alarms	Yes for all technologies
	Location accuracy of interpreted items	<0.2 m	Yes for all technologies
	Processing time (interpretation)	< 5 minutes operator time per anomaly	Yes for all technologies
	Accuracy of inversion parameters	Within class variance of cooperative inversion < single inversion	Yes for both MTADS EM61 and EM63

Table 4. Performance objectives for the Fort McClellan demonstration.

Type of Performance Objective	Primary Performance Criteria	Expected Performance (Metric)	Actual Performance (Objective Met?)
Quantitative	Survey rate	30 anomalies/day	No
	PDisc on recovered items at selected operating point	> 0.95	Yes
	False alarm rate with PDisc (recovered) = 0.95	> 50% reduction in false alarms	Yes
	False alarm rate with PDisc = 1	> 25% reduction in false alarms	Yes
	Location accuracy of interpreted items	<0.2m	Not applicable
	Depth accuracy of interpreted items	90% within 15 cm	Yes
	Accuracy of size parameter $L_1(t_1)$	Within class variation within one order of magnitude	Yes
	Accuracy of time decay parameter $L_1(t_{20})/ L_1(t_1)$	Within class variation within 25%	No
	Processing time (interpretation)	< 10 minutes operator time per anomaly	Not applicable
Qualitative	Reliability and robustness	Operator acceptance	No

5.0 SITE DESCRIPTION

Descriptions of each of the three sites visited are listed below.

5.1 FORMER LOWRY BOMBING AND GUNNERY RANGE

5.1.1 Site Location and History

FLBGR is located approximately 20 miles southeast of Denver, CO, in Arapahoe County. Although the area immediately west of the former bombing range is extensively developed, the site is still primarily grazing land. Evidence of DoD use of the bombing range remains at every known range. The gunnery ranges and small arms ranges still contain empty cartridges and projectiles.

FLBGR was originally part of Buckley Field, which consisted of the airfield and bombing and gunnery range and contained 65,547 acres. The status of the various portions of land that made up Buckley Field changed several times since the land was acquired by the City of Denver beginning in 1937. The airfield and bombing range were used by the Army during World War II. After the war, the airfield became a Naval Air Station, and the bombing range came under the custody of Lowry Air Force Base. The bombing range was renamed the Lowry Bombing and Gunnery Range. The bombing range was excessed beginning in 1960.

Within the demonstration areas there is little variation in terrain and vegetation. At both sites the vegetation is a mixture of grasses and yucca plants. These are dense, low-lying (< 1 m) plants that caused some survey difficulties to the EM63 cart in particular.

5.1.2 Munitions Contamination

In 2005, 45 acres on the RR, and 6 acres on the 20 mm RF were surveyed with the Sky EM61 towed array (Figure 6). These areas were specifically identified by the U.S. Army Corps of Engineers (USACE)-Omaha as priority areas that are currently being cleared (or will be cleared in the near future). The sites are also representative of the terrain, vegetation, and munitions at the site.

The RR was used for bombing practice with sand-filled practice bombs and high explosive (HE) bombs, rocket practice, and gunnery training. Expected UXO in this area include practice bomb debris, HE bomb fragments, 50-caliber ammunition and 20 mm projectiles and practice rockets. The 20 mm RF was used for air-to-ground target practice for fixed-wing aircraft firing 50-caliber projectiles, and 20 and 37 mm projectiles.

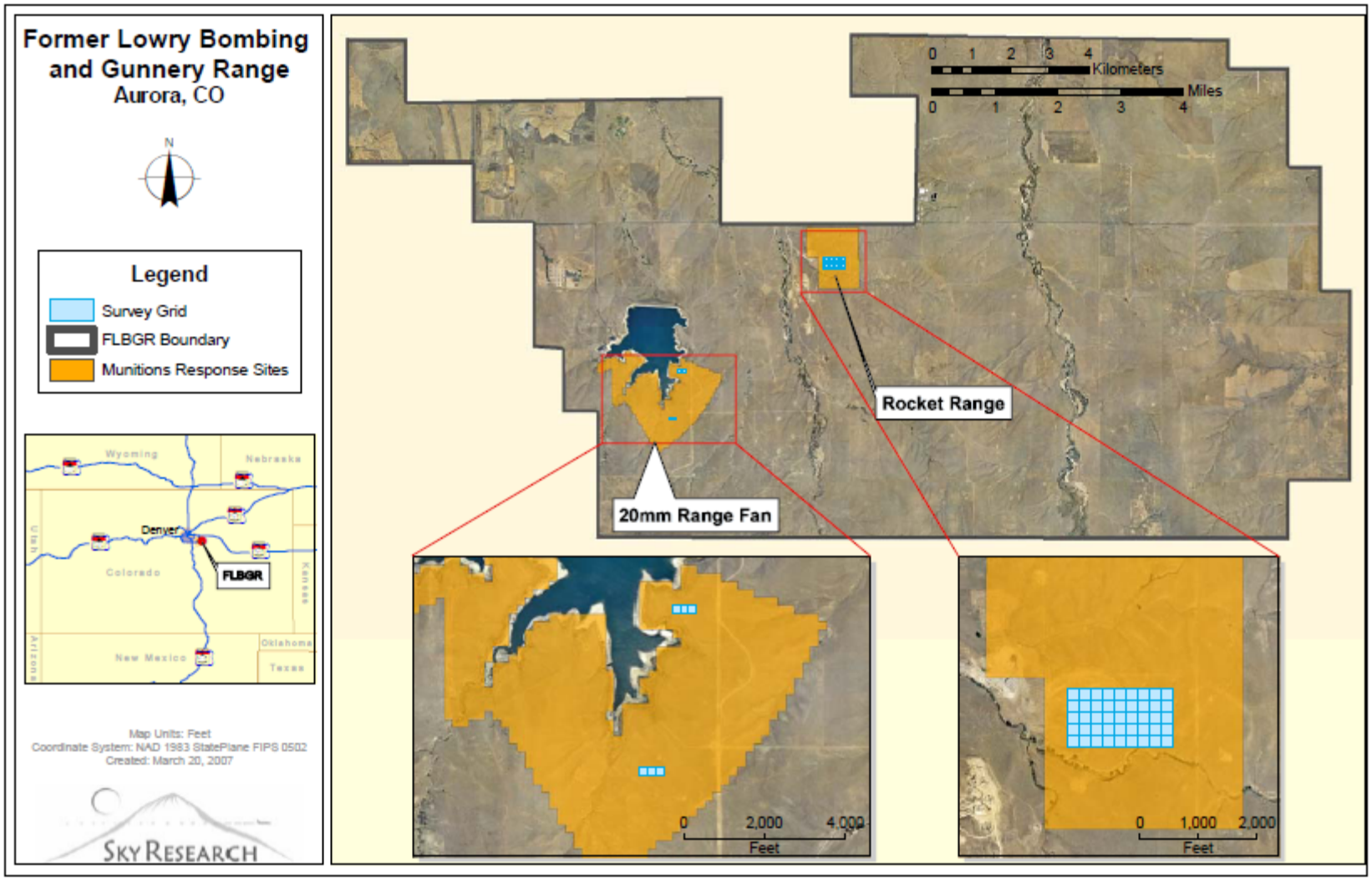

Figure 6. Locations of the RR and 20 mm RF sites at FLBGR.

5.2 FORMER CAMP SIBERT

5.2.1 Site Location and History

The Camp Sibert ESTCP UXO Discrimination Study Demonstration site is located within the boundaries of Site 18 of the former Camp Sibert Formerly Used Defense Site (FUDS). The land is under private ownership and is used as a hunting camp.

Information on the Camp Sibert FUDS is available in the archival literature such as an Archives Search Report (ASR) developed in 1993. The former Camp Sibert is located in the Canoe Creek Valley between Chandler Mountain and Red Mountain to the northwest, and Dunaway Mountain and Canoe Creek Mountain to the southeast. Camp Sibert consists mainly of sparsely inhabited farmland and woodland and encompasses approximately 37,035 acres. The City of Gadsden is growing towards the former camp boundaries from the north. The Gadsden Municipal Airport occupies the former Army airfield in the northern portion of the site.

The site is located approximately 50 miles northwest of the Birmingham Regional Airport or 86 miles southeast of the Huntsville International Airport. The site is near exit 181 off Interstate 59 in Gadsden and located approximately 8 miles southwest of the City of Gadsden, near the Gadsden Municipal Airport.

The area that would become Camp Sibert was selected in the spring of 1942 for use in the development of a Replacement Training Center (RTC) for the Army Chemical Warfare Service. The RTC was moved from Edgewood, MD to Alabama in the summer of 1942. In the fall of 1942, the Unit Training Center (UTC) was added as a second command. Units and individual replacements were trained in aspects of both basic military training and in the use of chemical weapons, decontamination procedures, and smoke operations from late 1942 to early 1945. Mustard, phosgene, and possibly other agents were used in the training. This facility provided a previously unavailable opportunity for large-scale training with chemical agent. Conventional weapons training was also conducted with several types and calibers fired, with the 4.2-inch mortar being the heavy weapon used most in training.

The U.S. Army also constructed an airfield for the simulation of chemical air attacks against troops. The camp was closed at the end of the war in 1945, and the chemical school transferred to Fort McClellan, AL. The U.S. Army Technical Escort Unit undertook several cleanup operations during 1947 and 1948; however, conventional ordnance may still exist in several locations. After decontamination of various ranges and toxic areas in 1948, the land was declared excess and transferred to private and local government ownership. A number of investigations have been conducted on various areas of the former Camp Sibert from 1990 to the present. These investigations included record searches, interviews, surface assessments, geophysical surveys, and intrusive activities.

The site is no longer in active use by the military. The demonstration area is owned by a single landowner who uses the area for a hunting camp. The discrimination study was conducted after the end of the hunting season between February and August 2007. Figure 7 is an aerial photo of the site with the discrimination study area survey locations shown.

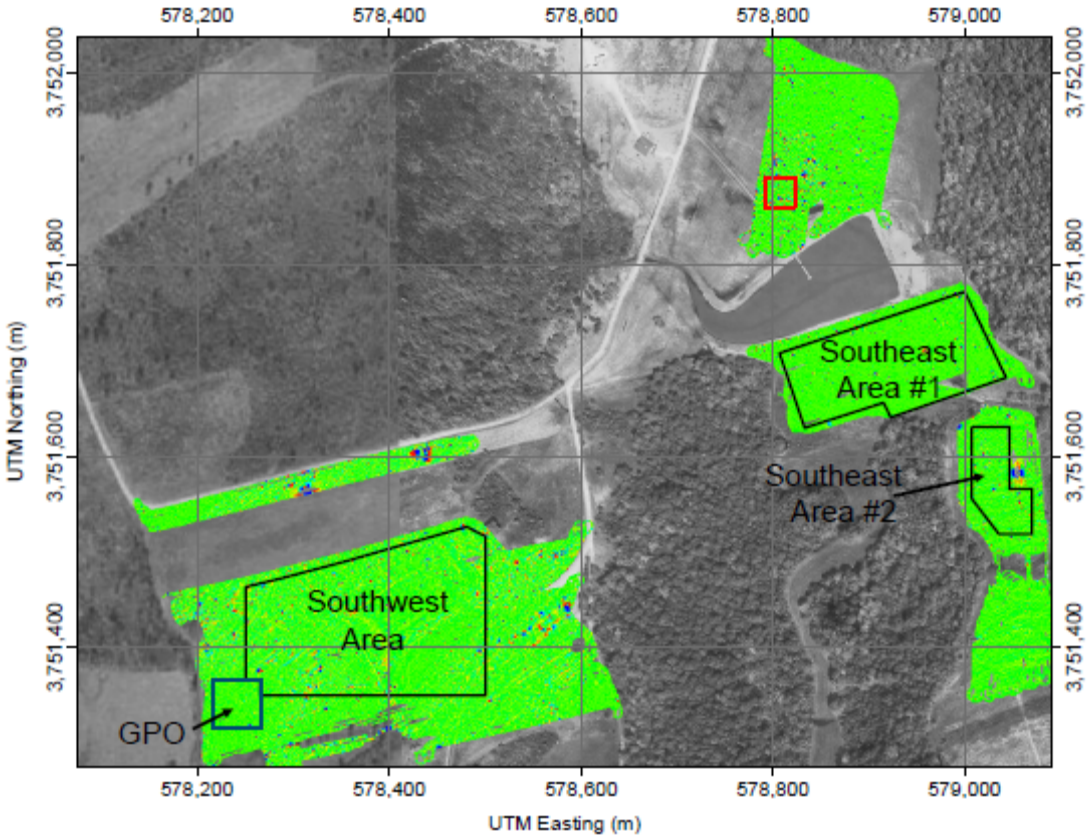

Figure 7. Camp Sibert site map with initial magnetometer survey locations shown.

5.2.2 Munitions Contamination

The ESTCP UXO Discrimination Study Demonstration Site is located within the confines of Site #18, Japanese Pillbox Area No. 2, of the former Camp Sibert FUDS. Simulated pillbox fortifications were attacked first with white phosphorus (WP) ammunition in the 4.2-inch chemical mortars followed by troop advance and another volley of HE-filled 4.2-inch mortars. Assault troops would then attack the pillboxes using machine guns, flamethrowers, and grenades. The locations of nine possible bunkers and one trench in 1943 were identified as part of a 1999 investigation. There is historical evidence of intact 4.2-inch mortars and 4.2-inch mortar debris at the site.

5.3 FORT MCCLELLAN

5.3.1 Site Location and History

The Fort McClellan test site was selected partly because the project could leverage the ongoing clearance activities being executed by Matrix Environmental and partly because it represented a physically challenging site to survey. Fort McClellan occupies 18,929 acres in the City of Anniston in Calhoun County, AL. To the west of Fort McClellan are the areas known as Weaver and Blue Mountain, and to the north is the City of Jacksonville. The Talladega Forest is located east of Fort McClellan. The portions of Fort McClellan to be addressed lie in the north-central portion of the installation, immediately adjacent to the main cantonment area. Figure 8 shows the

location of Fort McClellan and the four Munitions Response Sites (MRS) covered by this document.

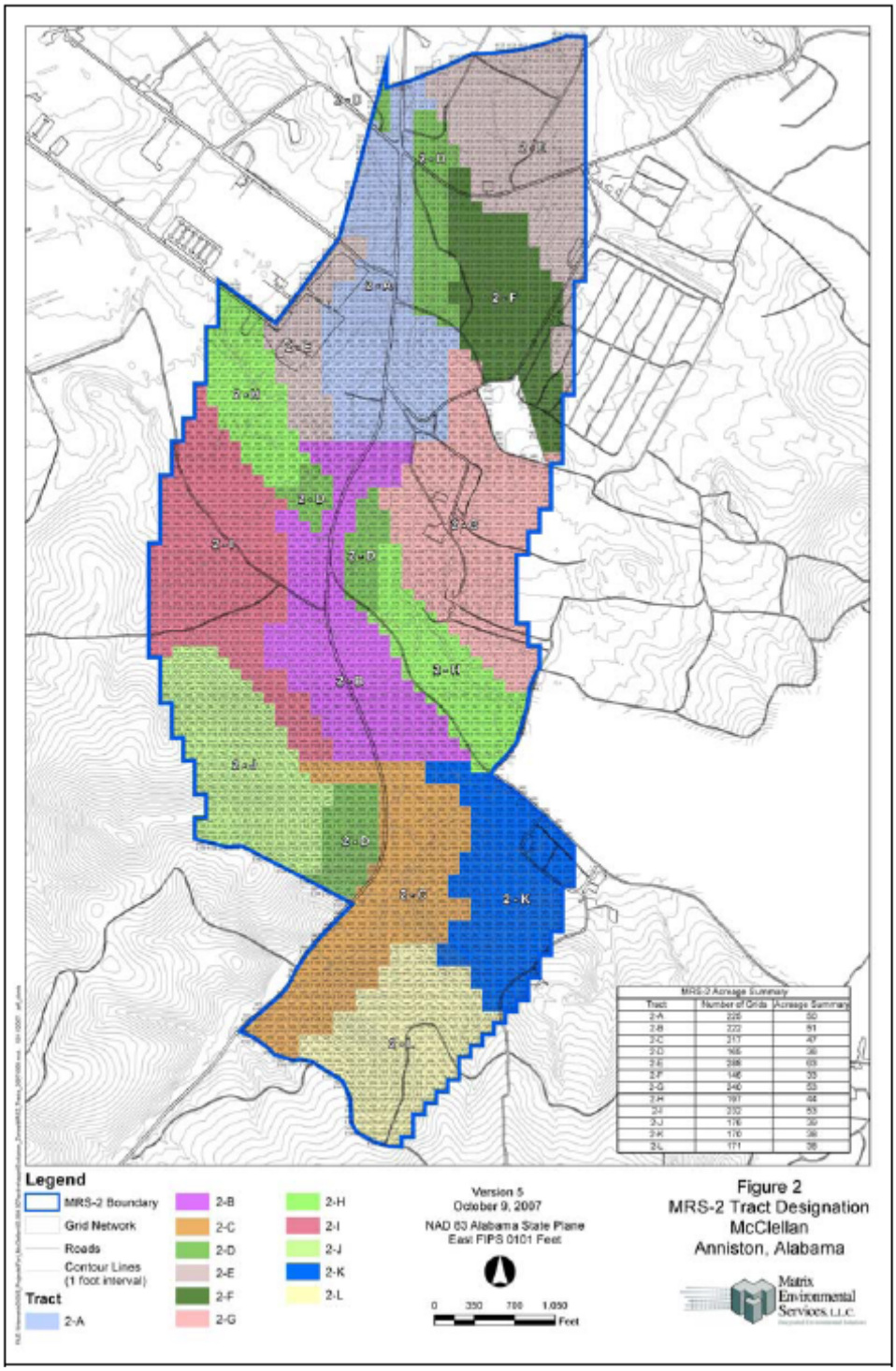

Figure 8. Fort McClellan site map.

Fort McClellan has documented use as a military training area since 1912, when the Alabama National Guard used it for artillery training. However, the Choccolocco Mountains may have

been used for artillery training by the units stationed at Camp Shipp in the Blue Mountain Area during the Spanish American War, as early as 1898. The 29th Infantry Division used areas of Fort McClellan for training prior to being ordered to France during World War I. In 1917, Congress authorized the establishment of Camp McClellan, and in 1929, the camp was officially designated as Fort McClellan. Prior to World War II, the 27th Infantry Division assembled at Fort McClellan for training, and during the war, many other units used the site for various training purposes. Following World War II, in June 1947, Fort McClellan was put in inactive status; it was reactivated in January 1950, and the site was used for National Guard training and was selected as the site for the Army's Chemical Corps School.

The history of Fort McClellan, includes training activities and demonstrations that used conventional weapons (i.e., mortars, anti-tank guns, and artillery pieces). Fort McClellan was recommended for closure under the 1995 Base Realignment and Closure Program and was officially closed in September of 1999.

The site is no longer in active use by the military. The Alpha Munitions Response Area (MRA) surrounds two active facilities, the Army's former Chemical Decontamination Training Facility (CDTF) and the Military Operations in Urbanized Terrain (MOUT). The CDTF is now referred to as the Chemical, Ordnance, Biological and Radiological Facility (COBRA) and has been transferred to the U.S. Department of Homeland Security. The MOUT is currently owned by the Alabama National Guard.

5.3.2 Site Geology

The Alpha and Bravo MRAs are predominantly heavily to moderately wooded with mixed pines and hardwoods, with some open areas that were cleared for various activities during the active operation of the installation. Numerous paved and unpaved secondary roads are present, along with occasional structures, many of which are no longer used.

Fort McClellan is situated near the southern terminus of the Appalachian Mountain chain. All but the easternmost portion of the former Main Post lie within the Valley and Ridge Province of the Appalachian Highlands. On a large scale, most of the rocks have been intensely folded into an aggregate of northeast-southwest trending anticlines and synclines with associated thrust faults. The shallow geology in the area is characterized by colluvial deposits. The presence of metamorphic rocks, as well as iron-bearing cements within the sedimentary rocks, increases the potential for minerals such as magnetite and other associated magnetic minerals.

5.3.3 Munitions Contamination

Provide a summary of what was known about the munitions contamination on the site. Include site maps illustrating the extent and distribution. Indicate the munitions types that were known or suspected to be present. The history of Ft. McClellan, includes training activities and demonstrations that used conventional weapons (i.e., mortars, anti-tank guns, and artillery pieces). 75 mm mortars and 3.8-inch shrapnel rounds were the primary munitions types considered likely targets of interest during data collection activities.

6.0 TEST DESIGN

6.1 FORMER LOWRY BOMBING AND GUNNERY RANGE

6.1.1 Conceptual Experimental Design

The specific objectives of the demonstration were to validate single and cooperative inversion approaches to UXO discrimination as a function of the following variables:

- COTS sensors
  - Single sensor, single data-type for inversion
    - Geonics EM61 as an industry standard COTS TEM sensor which provides four time gates at each sounding
    - Geonics EM63 as a higher quality COTS TEM sensor that provides 26 time gates spread over a larger range than the EM61
  - Dual sensor, dual data-type for cooperative inversion
    - Geonics EM61 and magnetometer
    - Geonics EM63 and magnetometer
- Type of munitions and clutter present
  - Eight grids in the RR at FLBGR contain a range of air-delivered munitions from 20 mm projectiles to large bombs.
  - Two grids in the 20 mm RF are primarily comprised of 50-caliber bullets and 20 mm and 37 mm projectiles. The ability to distinguish 50-caliber bullets and 20 mm projectiles (considered non-UXO) from 37 mm caliber projectiles (considered UXO) would be a significant advance at the site.
- Target density
  - Four grids in the RR and one in the 20 mm RF grid have high target density (> 150 targets per acre).
  - Four grids in the RR and one in the 20 mm RF have medium target density (50 to 150 targets per acre).
- Geological conditions
  - Two high density grids in the RR have soils that cause a measurable response in the EM61 data.

Each of the sensor systems used in the demonstration were positioned by a Leica RTS TPS 1206, with sensor orientation provided by a Crossbow AHRS400 IMU. The magnetometer data were of insufficient quality to be used for the cooperative inversion task and won't be considered further in this report.

The selection of grids for this demonstration was made by reference to previously collected EM61 towed array data. Magnetic data and Geonics EM63 data were collected over the 10

selected grids consisting of a total area of around 3.7 hectares (9.2 acres). Within the eight RR grids there were almost 1200 anomalies selected by reference to the towed-array data, while in the 20 mm RF there were 407 anomalies. Figure 9 and Figure 10 show the grids surveyed within the RR and 20 mm RF.

6.1.2 Site Preparation

The majority of mobilization activities for this demonstration were completed as part of the concurrent geophysical surveys being conducted for Army ERDC and USACE-Omaha by Sky. Surface clearance was already conducted and survey control established. Additionally, Sky has an on-site trailer and storage compound within a few kilometers of the RR site. As noted previously, the EM61 data were collected during a previous ERDC sponsored mobilization in September and October 2005.

Project-specific mobilization consisted of the following:

1. Mobilization of the EM and magnetometer field crew and associated equipment.
2. Mobilization of the quality assurance officer to the site.
3. Emplacement of twenty 37 mm projectiles within the 20 mm RF. This provided training data for the EM63 and magnetometer sensors and also served as a test of the detection performance of the sensor systems (against the smallest UXO).
4. Standard pre-collection maintenance and calibration procedures were performed for the sensor systems. These included all the calibrations listed in our Quality Assurance Project Plan in Appendix C of the Demonstration Plan.

6.1.3 System Specification

Three different systems were deployed at FLBGR, including a magnetometer, an EM61 array and an EM63.

Sky's EM61MK2 towed array (Figure 11) contains five coils, Crossbow AHRS400 IMU, and Leica RTS. Data at the RR and 20 m RF were previously collected by this system. The EM61 logged data at 10 Hz, the RTS at 4 Hz, and the Crossbow IMU at 30 Hz. Sky's modified EM63 cart system with Leica RTS for position and Crossbow IMU for sensor orientation was used to conduct the second survey of the area (Figure 12). EM63 data at 26 geometrically spaced time gates (spanning the range 180 μs to 25.14 milliseconds [ms]) were collected at a 5 Hz rate (the maximum for the EM63). The RTS was operated at around 4 Hz and the Crossbow IMU at 30 Hz. The EM63 coil was 25 cm above the ground and was used to collect data along transects spaced 0.5 m apart. Data were collected while walking slowly at about 2 km/hr so that the along line sample spacing was approximately 10 cm.

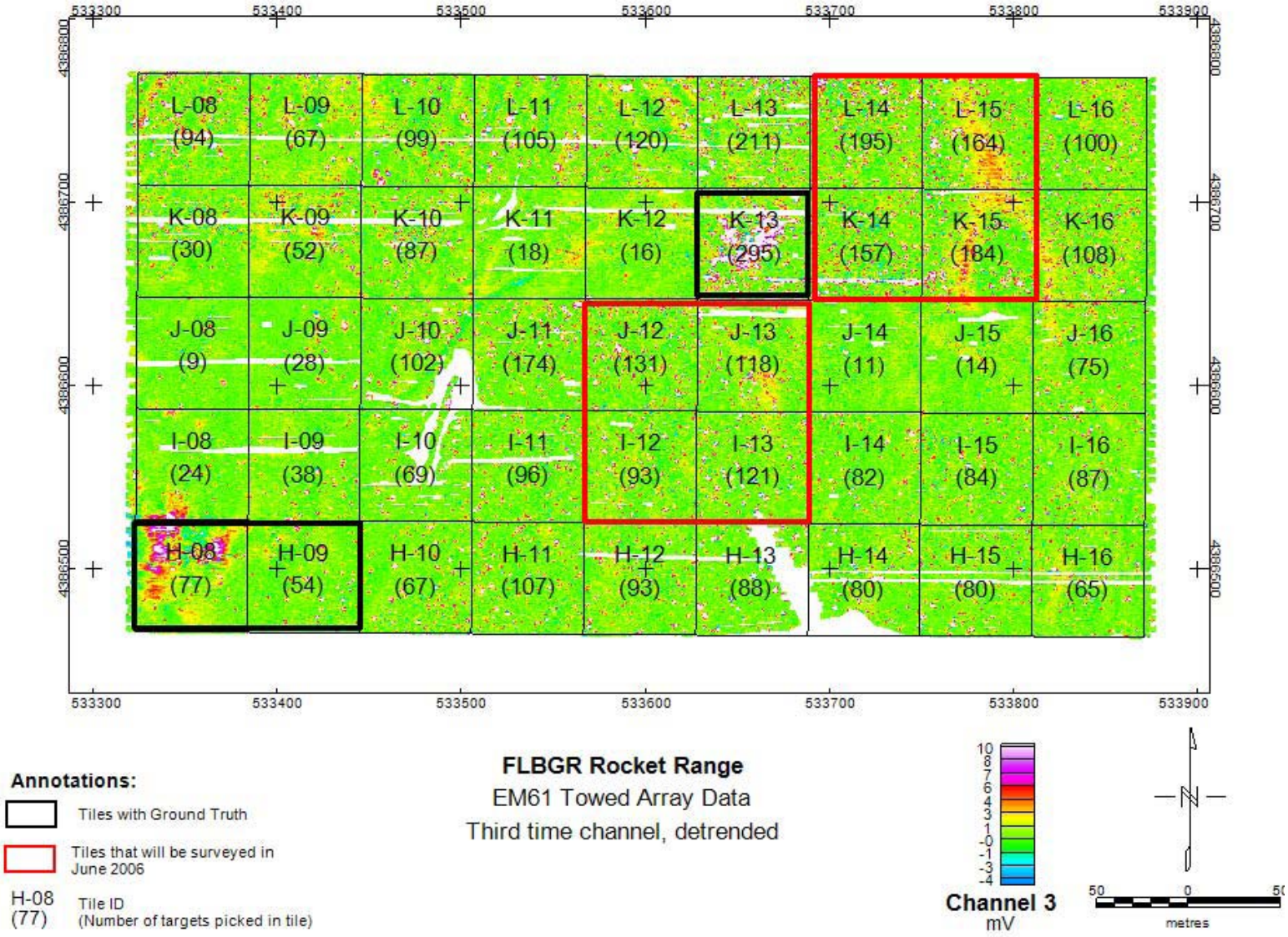

Figure 9. Map of RR with areas surveyed for this demonstration outlined in red.

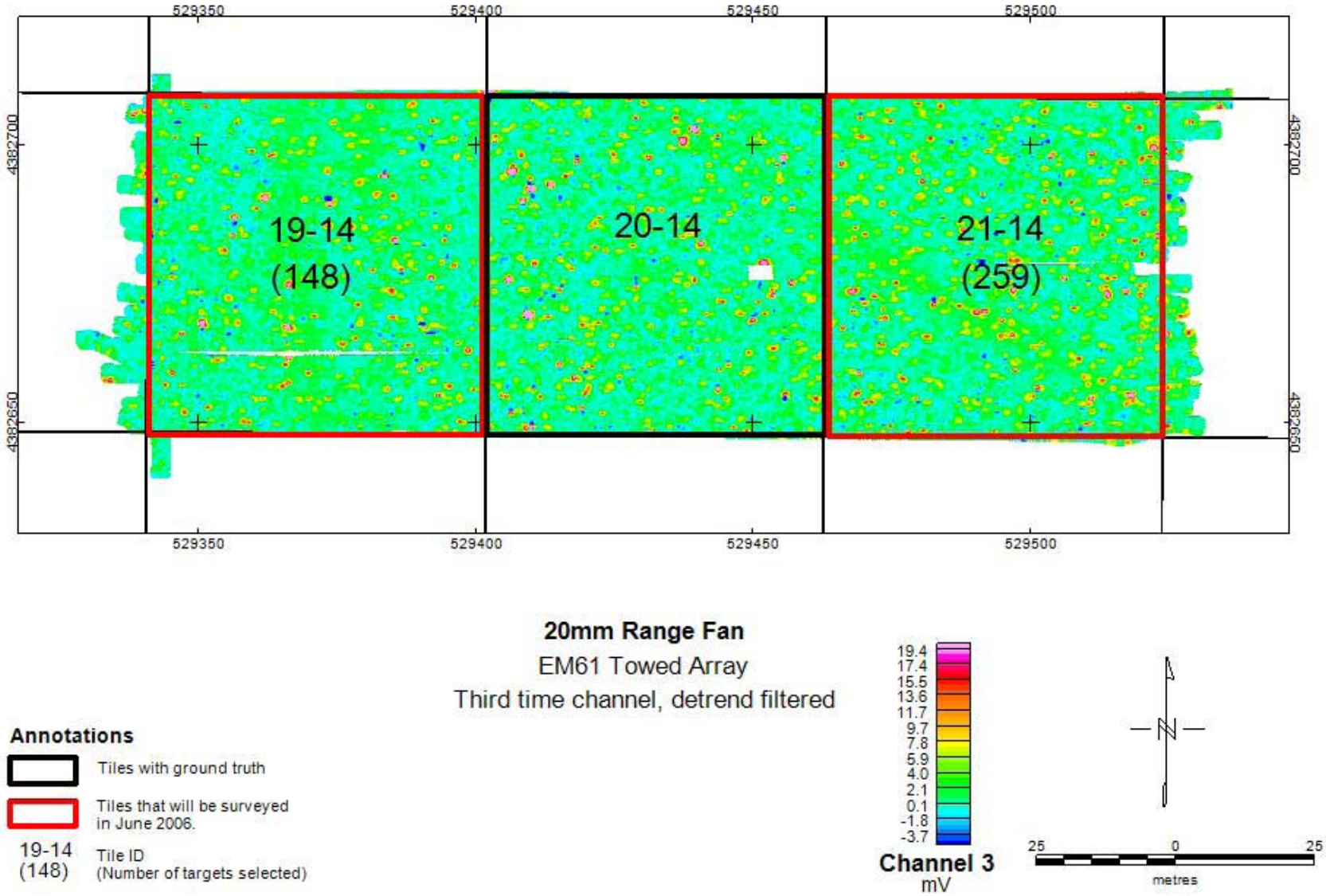

Figure 10. Map of the 20 mm RF, with two grids surveyed for this demonstration outlined in red.

Figure 11. Sky's EM 61-MK2 towed array, which is constructed of composite materials and houses EM sensors, RTS laser positioning (or GPS) sensors, and the Crossbow IMU.

Figure 12. Equipment used at FLBGR including the modified EM63 cart (left) and the Leica RTS TPS1206 laser positioning system (right).
This device is set up in over a known point and tracks a prism attached to the geophysical survey equipment.

6.1.4 Data Collection

Table 5 lists the project's key activities.

Table 5. Key project activities.

Day	Activity
Pre-Survey	
Sep-05	Commence EM61 data collection over 45 acres on RR, and 6 acres on the 20 mm RF
Oct-05	Complete EM61 data collection and initial processing
Jan-06	Validation of five grids (4 on RR, 1 on 20 mm RF)
May-06	Demonstration Plan approved

Table 5. Key project activities. (continued)

Day	Activity
EM63	
12-Jun-06	Arrive on site
13-Jun-06	Commence EM63 survey of J-13
14-Jun-06	Continue with J-13
15-Jun-06	Complete J-13, start J-12
19-Jun-06	Survey test plot with EM63
20-Jun-06	Troubleshoot problems with CSM EM63. Down-time while waiting for replacement instrument
23-Jun-06	Receive replacement EM63 and survey test plot
26-Jun-06	Start surveying J-12 with new EM63 (repeat areas down with the CSM system)
27-Jun-06	Finish J-12, start and finish I12
28-Jun-06	Recollects on I12, start I-13. Strong winds knock over the RTS base-station and damages it
29-Jun-06	Receive replacement RTS and continue with survey of I-13
30-Jun-06	Demob from site for 1 week (short week due to 4th of July celebration)
9-Jul-06	Redeploy to site
10-Jul-06	Commence EM63 survey of 19-14
11-Jul-06	Complete survey of 19-14, commence 21-14
12-Jul-06	Complete 21-14, recollects on 19-14, move equipment back to RR
13-Jul-06	Start and complete K-15, commence L-15
17-Jul-06	Complete L-15
18-Jul-06	Start and complete L-14
19-Jul-06	Start and complete L-13 and conclude ESTCP surveys
Processing	
Jun-06	Initial processing EM63
Jul-06	Complete initial processing EM63
Aug-06	Feature extraction commences
Validation	
Aug-06	Delay excavations to wait for a possible MSEMS survey
11-Sep-06	Submit Phase I interpretations (I12, J-12, 19-14)
11-Sep-06	Excavation commences of Phase I grids
13-Sep-06	Commence excavation of Phase II grids on 20mm RF (21-14)
14-Sep-06	Phase I ground truth released
18-Sep-06	Phase II interpretation on 20 mm RF (21-14) submitted
19-Sep-06	Phase II ground truth released for 20 mm RF (21-14); Phase II validation of RR commences (L-15)
19-Sep-06	Phase II validation of K-15 commences
21-Sep-06	Phase III validation commences (I-13, J-12)
22-Sep-06	Phase II interpretation on RR (K-15, L-15) submitted
25-Sep-06	Phase III validation continues (L-13, L-14)
5-Oct-06	Phase III interpretations submitted (I-13, J-12, L-13, L-14). Later that day, Phase III ground truth released

*MSEMS=Maintenance Standardization and Evaluation Management System

6.1.5 Validation

Ground truth data were managed by Linda Daehn of Tetra-Tech EMI Inc. of Helena, MT, and were kept secret from Sky. Ground truth was only released after interpretations had been submitted to the ESTCP Program Office. There were two phases of ground truth data release on the 20 mm RF and three phases on the RR. In Table 6 we list the total number of items recovered in each grid in each of the following categories MK-23, 2.25-inch rocket, 37 mm projectile, 20

mm projectile, small arms (50-caliber), shrapnel, and junk. We consider MK-23s and 37 mms as UXO regardless of whether they are inert or live rounds.

Table 6. Number of anomalies identified as MK-23, 2.25-inch rocket, 37 mm projectile, 20 mm projectile, small arms, shrapnel, and junk in each of the 10 grids used for this demonstration.
For each grid, we identify the phase where ground truth was released.

Grid and (Phase)	Bomb Mk-23	Rocket 2.25-inch	37-mm Projectile	20-mm Projectile	Small Arms	Shrapnel	Junk	Grand Total
19-14 (I)			18	25	91			**134**
21-14 (II)			20	39	90		1	**151**
I-12 (I)	5			46	4	28	2	**85**
I-13 (III)	14			41	8	20	10	**93**
J-12 (I)	7			44	14	52	17	**134**
J-13 (III)	3			45	9	34	10	**101**
K-15 (II)				52	27	44	60	**183**
L-13 (III)		1		55	23	72	44	**195**
L-14 (III)				65	14	66	51	**196**
L-15 (II)				50	18	54	43	**165**
Grand Total	**29**	**1**	**38**	**462**	**298**	**370**	**238**	**1437**

On the 20 mm RF, the majority of recovered items were either 50-caliber bullets or 20 mm or 37 mm projectiles. There was almost no shrapnel or junk recovered from that site. Of the 37 mm projectiles, the majority were seed items emplaced by Sky in known locations (20 rounds) or by ERDC (16 items) in unknown locations.

On the RR, there are also large numbers of 50-caliber bullets (small-arms) and 20 mm projectiles, but no 37 mm projectiles. All but one UXO were MK-23 practice bombs, with a single rocket warhead recovered. In addition, almost all the MK-23s were found at or near the surface so they had large amplitudes and large SNR. Unfortunately, for the two Phase II grids (L-15 and K-15) and two of the Phase III grids (L-13 and L-14), there was only one UXO item found (the rocket warhead), and the fit for that item was failed. The number of items on the dig-sheet is the only metric we can use to compare the amplitude and classification methods. The best method will be the one that recommends the lowest number of items to be excavated regardless of the validity of the underlying discrimination methodology. Therefore, results for Phase II RR grids and L-13 and L-14 in Phase III were not used to compare the performance of the discrimination methods.

6.2 FORMER CAMP SIBERT

6.2.1 Conceptual Experimental Design

At Camp Sibert, we used data collected using various platforms by different demonstrators. The specific data modeling activities conducted were:

- Dipole fitting of the MTADS magnetometer data and calculation of the magnetic remanence metric.

- Fitting of 3-dipole beta models to the MTADS EM61 data. This hybrid model fitting approach is an attempt to prevent UXO being incorrectly modeled with three distinct polarizations, while allowing the model enough flexibility to model irregularly shaped shrapnel.
- Fitting of 3-dipole beta models to the contractor EM61 data.
- Fitting of 3-dipole Pasion-Oldenburg models to the EM63 cued-interrogation data.
- Cooperative inversion of the EM61 and EM63 data using the dipole fits from the magnetometer data to constrain the object's location and depth.

For interpretation we submitted two different types of dig sheets. The first was based on size parameters alone, while the second used all parameters and statistical classification. The following four size-based interpretations were delivered:

1. *Magnetics size-based (moment).* A dig sheet ranked according to decreasing size of the recovered dipole moment
2. *Magnetics size-based (remanence).* A dig-sheet ranked according to magnetic remanence calculated using models designed to represent the induced magnetization of the items expected at the site
3. *MTADS EM61 size-based.* Production of a dig sheet ranked according to size (using the sum of the beta parameters for either time channel [1]).
4. *Contractor EM61 size-based.* The same as (2) but with the contractor data set.

For the statistical classification, we used data over the GPO and the initial training grids to determine the feature vectors and statistical classifier to use. The following statistically based dig sheets were produced:

1. *MTADS EM61 statistical.* Using statistical classification of features derived from the MTADS EM61 data
2. *Contractor EM61 statistical.* Using statistical classification of features derived from the contractor's EM61 data
3. *EM63 statistical.* Same as 2 but with the EM63
4. *MTADS EM61 and magnetics statistical.* As per 2 but with MTADS EM61 fits constrained by the magnetics data and with the addition of the features from the magnetometer data (remanence, moment, etc.)
5. *EM63 and magnetics statistical.* As per 4 but with the EM63.

6.2.2 Site Preparation

An initial magnetometer survey of the Camp Sibert site was used to select three different areas to be used in the discrimination study: South-East-1 (SE1), South-East-2 (SE2) and South-West (SW). The three areas combined covered approximately 15 acres. A GPO was established

immediately adjacent to the SW area and consisted of thirty 4.2-inch mortars emplaced at different depths and orientations and eight "partial" mortars that are considered nonhazardous scrap. ESTCP emplaced 152 4.2-inch mortars in the SE1, SE2, and SW areas. The locations and depths of 29 of these mortars, along with the locations and identities of 179 "clutter" items were provided to each demonstrator and used as training data.

6.2.3 System Specification and Data Collection

In full coverage survey mode, data were collected using the standard cart platform of the EM61 MK2 system. EM61 MK2 cart data were acquired by an on-site contractor using a line spacing of 50 cm, sensor height of 40 cm, and positions recorded with a Real-Time-Kinematic Global Position System (RTK-GPS), accurate to within a few centimeters. Survey mode data were also acquired for both a magnetometer array as well as an EM61 MK2 array using the MTADS. The MTADS EM consists of three overlapping EM61 MK2 sensors (each 1 m wide and 0.5 m long) that have a center-to-center separation of 0.5 m (Nelson et al., 2003). Data are collected with a nominal across track sensor spacing of 50 cm, and the array is transported 25 cm above the ground. The MTADS magnetometer array consists of a platform housing a 2 m wide in-line array of eight G-822 Cs vapor magnetometers spaced 25 cm apart and 25 cm above the ground (Nelson et al., 2003). For both the MTADS EM and magnetometer arrays, position and orientation information were obtained through centimeter-level RTK-GPS and IMU measurements.

The EM63 was deployed in a cued-interrogation mode on a customized air suspension cart. The cart served both to lower the instrument closer to the surface (from the standard 40 cm to 20 cm) as well as to absorb some of the effects from instrument jostle due to an uneven ground surface. The late time information available from the 26 time channel EM63 provides an extended view of target decays. Positional information was collected via a Leica TPS 1206 RTS with orientation effects recorded using a Crossbow AHRS 400 IMU. We estimate that positions were accurate to within 2-4 cm and orientation to within about 2°.

6.3 FORT MCCLELLAN

6.3.1 Conceptual Experimental Design

The objectives of this demonstration were to evaluate the discrimination potential of the Geonics EM63 at Fort McClellan, AL, when deployed in a cued interrogation mode. Pasion-Oldenburg polarization tensor models were fit to each of the EM63 cued anomalies. Feature vectors extracted from those dipole fits were used to guide a statistical classification algorithm that ranked the items in order of UXO likelihood.

6.3.2 Site Preparation

The survey areas were cleared of brush and any freestanding trees with a diameter of less than 3 inches prior to the geophysical field efforts. Even with this clearance, there were substantial mature trees throughout the survey area that prevented the use of either GPS or laser theodolite positioning systems. The areas were initially mapped with EM61 full coverage surveys performed by NAEVA Geophysics, Inc. (NAEVA) (e.g., Figure 7). Positioning was achieved

using tape measures and ropes set out from staked corner locations of 50 ft x 50 ft grids. Target picks were made by NAEVA personnel, and an EM61 reacquisition team was deployed to confirm the location of the picked anomalies and plant labeled flags at the respective reacquired target locations.

6.3.3 System Specification

The system deployed at Camp Sibert included the EM63, an orientation sensor and a survey template. We used the Geomechanics MD900-TS Digital/Analog Clinometer (http://www.geomechanics.com/pdf/products/MD900T%20IRIS,%20L00251C.pdf) with viscous damped sensor. This system has a measurement range of 25°in pitch and roll, with a resolution of better than 0.004° and repeatability within 0.2°.

The cued interrogation procedure consists of surveying a 2.5 m x 2.5 m area over pre-identified locations in a star pattern, as illustrated in Figure 13. The survey template was always oriented so that first point and the corresponding 2.5 m long line, were oriented from West to East (unless a tree or obstacle was in the way). In order to gauge instrument drift and obtain a measure of background geology, data were collected on the four corners of the template both before and after surveying. This usually placed the EM63 far enough away from the anomaly, but there were a few cases where a nearby anomaly contaminated the result. By collecting four points, we had enough redundancy so that at least two points sampled the background geology.

Figure 13. Standard EM63 cart collecting discrimination mode data at the Ashland test site.

A rigid fiberglass indicator rod is mounted in the center of the coil to accurately log the survey location. Along with the inclusion of the IMU for orientation, this minimizes positional errors.

Predeployment testing in Ashland indicated that a thicker gauge plastic was preferable to a tarp as it was less likely to bunch up or shift during surveying. The weight of the mat is also substantial enough that it will not be easily windblown or require any ground intrusive means to be secured. Data were collected by pushing the EM63 so that it was centered directly over

marked locations on the survey mat and held static for 2 second recording intervals. In order to ensure that the cart was centered over the indicated point on the template, a semi-rigid fiberglass rod was positioned such that it extended from the center point of the coils to just above the surface.

6.3.4 Data Collection

The field team deployed to Anniston on March 16, 2008, and arrived onsite at Fort McClellan on March 17. An inventory of the EM63 shipment that had arrived the previous week was taken after receiving a brief tour of the site from Kent Boler of Matrix. Unfortunately, FedEx had lost part of the equipment shipment, including the cart wheels, console, batteries, and some cables. A replacement EM63 was sent from the Colorado School of Mines and arrived early on the afternoon of March 17. The equipment was assembled and some simple tests were run in the parking lot to confirm that equipment was operational. On March 18, a formal site-induction took place with the NAEVA field personnel and surveying commenced immediately afterwards. The following general procedures were followed for the survey:

- After collection of the start-of-day calibrations, the sensor operator verified correct operation of all system components prior to commencing data collection.
- For each anomaly, the third crew member set-up the template over the flagged location in preparation for the arrival of the EM63 and its two operators. The anomaly was then surveyed.
- The four corners of the survey template were surveyed before and after collecting the 55-point pattern on the template. The corners were collected to provide a static background measurement with the EM63 at a nearby, source-free location.
- The third crew member then packed up the template and moved it to the next anomaly. We used two templates so that the EM63 was in continual use.

It became evident after the first few full days of surveying that the flagged locations were not always consistent with the peak of the observed EM63 response, leading to some anomalies that were poorly centered. Because the template was designed with points increasingly clustered towards the center of the template, collecting the highest quality data required the template to be accurately centered over the intended target response. To ensure centered targets, the template was placed directly over the flagged location and the EM63 was then used to scan the areas immediately surrounding the center point to ensure that the target was centered. When the maximum EM63 response did not correspond with the flagged location, the template was shifted to ensure that the maximum target response was centered underneath the survey template. A 2.5 m by 2.5 m section of data were collected around 401 anomalies identified by Matrix Environmental along with:

- Eighteen targets from the GPO
- Seven typical targets from the site at three unique orientations (horizontal, vertical, 45°) for a total of 21 additional cued interrogation surveys, plus two surveys with no item to provide an estimate of the background noise.

Table 7 lists the number of anomalies surveyed in each grid, along with a count of the different types of items in each grid. Figure 14 shows some pictures illustrating data collection challenges encountered at the site.

Table 7. The number of items surveyed in each grid.

Grid	Cultural Debris	MEC* Scrap	Medium MEC	No Find	Small MEC	Small Arms	Small-Medium MEC	Total
N069E143		17						17
N069E144	2	12						14
N070E143		6	9					15
N071E141	1	7					1	9
N071E142	1	9	2					12
N071E143		12	4					16
N071E144		12	1					13
N071E145		5						5
N071E146	2	4						6
N071E147		2						2
N071E148		3						3
N071E149	1	1						2
N071E150		2	1			1	1	5
N071E152		2						2
N072E141	2	5						7
N072E142	2	6						8
N072E143		16						16
N072E144	1	10	4	1				16
N072E145		2	1					3
N072E146		6	1					7
N072E147	1	4	1					6
N072E148	1	2	1					4
N072E149		1						1
N072E150		2						2
N072E151	2	2						4
N073E141	1		2					3
N073E142		5	5					10
N073E143			12					12
N073E144	1	4	9					14
N073E148	2	1				1		4
N073E149		2				2		4
N073E150		2						2
N073E151		5						5
N074E141	1	1	1					3
N074E142	1	3	2					6
N074E143	1	3	4	1				9
N074E144	3	1	5					9
N074E148	1	2						3
N074E149	1	5	1					7
N074E150		2						2
N077E143	1	3						4
N077E144	1	5						6

Table 7. Count of the number of items surveyed in each grid. (continued)

Grid	Cultural Debris	MEC Scrap	Medium MEC	No Find	Small MEC	Small-Arms	Small-Medium MEC	Total
N077E145		2	2					4
N077E146	5	1	1					7
N077E147	4	3	1					8
N078E143	1	3	2					6
N078E144	1	1	2			1	1	6
N078E145	1	4	2					7
N078E146	3		1					4
N078E147	1		1					2
N079E143		4	1					5
N079E144	3	3	4					10
N080E143	1	10		1		1		13
N080E144	1	5						6
N081E144	1	2						3
N082E144	3	5	1			1		10
N082E145	2	8	1				1	12
Test Pit			9	2	3		9	23
GPO			5		2		11	18
Grand Total	57	245	99	5	5	7	24	442

*MEC = munitions and explosives of concern

Figure 14. Surveying difficulties encountered on the site including (a) steep slopes, (b) rough surfaces, and (c) cramped survey areas due to closely spaced trees.

6.3.5 Validation

Validation sheets compiled by the explosive ordnance disposal (EOD) technicians after anomaly excavation were the primary source of information used for performance confirmation. During anomaly validation, the EOD technicians (who worked under the direction of Matrix Environmental) recorded (1) the anomaly source (UXO and type, shrapnel, junk, no-find); (2) depth of burial to the top of the item; (3) azimuth and dip (for UXO items); and (4) approximate weight (for shrapnel and junk). The dig team also photographed any items that required demolition. We had requested that the dig team record the bearing and distance from the flag, but this information was not recorded.

7.0 DATA ANALYSIS AND PRODUCTS

7.1 PREPROCESSING

There were slight variations in the specific preprocessing steps applied to each of the datasets at each of the demonstration sites but they all broadly followed these steps:

- *Initial review of collected data.* Confirm that data fall within prescribed recording ranges, establish number of points collected, data density, and time-on/time-off. Reject invalid readings in either the sensor, positional, or orientation streams.
- *Data merging.* The sensor, positional and orientation data are merged using a common time reference to produce a data stream with positions, orientations, and data measurements at each observation location.
- *Drift and background correction (FLBGR).* Sensor drift and background variations were removed from the sensor data by subtracting a moving demedian value from each data channel. For the EM61 and EM63, the window length consisted of a fixed time between 10 and 15 seconds. For the magnetometer data, the window length was a fixed spatial distance between 10 and 15 m.
- *Drift correction (at Fort McClellan).* Repeat measurements were made on the four corners of the template with any difference in the repeat measurements attributed to sensor drift. For each recorded time gate, a drift correction is applied that was a linear interpolation (as a function of time) between the average of the before and after repeat measurements.
- *Background removal (Fort McCellan).* An estimate of the soil background (assumed constant over the breadth of the template) is calculated for each time channel using the median value of the lowest 50% of the measurements (to avoid biasing the background estimation with signal from the metallic anomaly).
- *Data gridding.* Filtered data are interpolated onto a 0.1m grid and reviewed by a geophysicist.

7.2 TARGET SELECTION FOR DETECTION

At FLBGR, an automatic target picking algorithm was applied to the EM61 and EM63 data. A threshold of 10 millivolts (mV) on time channel 3 was used for the picking of EM61 anomalies, and a threshold of 35 mV on time channel 1 for the EM63. The picking threshold was chosen by reference to the data collected over the FLBGR test plot.

At Camp Sibert, detection of anomalous regions was conducted by the demonstrators who provided the easting and northing coordinates corresponding to the maximum geophysical signature for each anomaly. The demonstrators used an automatic target detection algorithm that triggered a detection each time the sensor response exceeded a threshold. The threshold value was set to one-half of the minimum amplitude recorded over any of the 4.2-inch mortars in the GPO. At Fort McClellan, the EM63 data collections were cued from EM61 data previously collected by NAEVA. They provided anomaly lists and maps prior to deployment to the site. The

maps included an image of the EM61 response, anomaly picks (based on a 7 mV threshold), and the approximate locations of trees. We manually reviewed the target lists and maps and removed any anomalies that overlapped other anomalies or that were close to trees or other obstacles. The list of suitable anomalies was significantly larger than our survey goal of 400 items. To reduce the list to 400 items, we used random selection. We first created a histogram of the anomaly amplitudes and then subjectively split the anomalies into low, medium, and high SNR. Random selection was used to select about 133 anomalies from each of these categories.

7.3 PARAMETER ESTIMATES

Parameter estimation procedures for the three demonstrations were similar, except that at each successive demonstration, the procedures and algorithms improved. The following steps were used to extract features over each anomaly:

- *Formation of covariance matrix (EM61 and EM63).* The base-level noise in the data (as a function of both space and time) was determined using an automated procedure, and that noise floor was assigned to each anomaly. Together with a percentage error term, the base-line error was used to form the data-covariance matrix.
- *Region definition.* For each picked anomaly, a region of data for submission to the inversion algorithm was automatically selected. Where necessary, this automated selection was manually modified by an analyst.
- *Single inversion (magnetics).* A static dipole together with a constant shift was fit to each anomaly.
- *Single inversion (EM61).* Three-dipole instantaneous polarization models were fit to each EM61 anomaly.
- *Single inversion (EM63).* Pasion-Oldenburg parameterized three-dipole models were fit to each anomaly.
- *Cooperative inversion (EM61 and EM63).* The position and depth of the static magnetic dipole were used as constraints in the EM61 and EM63 inversions.

In all datasets, there are a number of targets where the observed fit obtained through the inversion process is unsatisfactory and must be failed. These failed fits were classified as can't-analyze and would have resulted in a significant number of excavations. We therefore investigated methods to reduce the number of items that need to be excavated while hopefully maintaining the same probability of correct classification.

Close scrutiny of failed inversions and inversions that yielded inaccurate depth estimates on the GPO and ground truth targets revealed that poor inversions could be tied to certain features of the data or the inversion. This motivated us to identify and establish rules to define a confidence factor for a given inversion. We considered several criteria that we gathered under a so-called Figure of Merit (FOM), which consists of:

- Data features
  - SNR. SNR should be above a given threshold for reliable inversion of each time channel. SNR should decay with time if the sensor operates properly and noise estimates are accurate.
  - Data coverage of anomaly. Coverage should sample the spatial decay of the EM scattered field to allow recovery of orthogonal polarizations.
- Inversion features
  - Quality of fit. Misfit, correlation coefficient
  - Variance of estimated depth. There can be several solutions of the inverse problem with similar misfits but distributed over a large range of depth.

Full details and quantitative descriptions of each of the FOM metrics are provided in Lhomme et al. (2008).

7.4 CLASSIFIER AND TRAINING

The following general procedure was used for the feature vectors of each sensor combination:

- *Selection of features*. By analysis of the training data, those features that contribute to separation of the different classes (consisting of UXO types and clutter) were selected. For both the EM61 and EM63, a combination of a size and time-decay feature vector provided reliable classification ability.
- *Choice of classification algorithm*. Through analysis of the training data, the best performing classifiers were selected: a Probablistic Neural Network.
- *Classification*. Anomalies were placed on a prioritized dig list by using the classifier to compute probabilities of class membership for unlabeled feature vectors. The probability of membership of the UXO class was reported on the dig sheet.
- *Anomalies where feature vectors are unreliable*. Some anomalies had insufficient SNR or data coverage to constrain the TEM model parameters. This included anomalies with overlapping signatures that could not be isolated and inverted one at a time. All these anomalies were placed in the dig sheet and labeled can't-analyze and were excavated as suspected UXO.

More details on the classification methods used are provided in the performance assessment section.

7.4.1 Classification Strategy at FLBGR

The objective in the 20 mm RF was to discriminate between ubiquitous small UXO (20 mm projectiles) and larger 37 mm projectiles. A number of 37 mm items were emplaced by ERDC in surveyed grids; ground truth was available for some emplaced items, but not all. In the RR, a wider variety of UXO were expected, with items of interest ranging in size from 100 lb bombs

down to 37 mm projectiles. A training data set for both areas was obtained from geophysical prove-out (GPO) data acquired over the FLBGR test plot that was established several years previously to support production geophysical activities. Emplaced items in the test plot were representative of all types of UXO expected in the two survey areas.

Data from two grids within the RR (one medium, one high density) and one grid within the 20 mm RF (medium density) were interpreted using the training data from the test plot, and ranked dig lists were submitted to the ESTCP Program Office and Don Yule of USACE-ERDC, together with a recommendation for how many of these anomalies to dig as potential UXO. After ground truth for these grids was collected (for all detected anomalies, not just the ones recommended for excavation) they were released to the analysts who then updated the classification strategy for the remaining grids. Complete ground truth from two more of the RR grids (medium and high density) was released and revised dig sheets submitted for the remaining four grids.

On the 20 mm RF, the majority of recovered items were either 50-caliber bullets or 20 mm or 37 mm projectiles. There was almost no shrapnel or junk recovered from that site. Of the 37 mm projectiles, the majority were seed items emplaced by Sky in known locations (20 rounds) or by ERDC (16 items) in unknown locations. On the RR, there were also large numbers of 50-caliber bullets (small arms) and 20 mm projectiles, but no 37 mm projectiles. All but one UXO were MK-23 practice bombs, with a single rocket warhead recovered. In addition, almost all the MK-23s were found at or near the surface so they had large amplitudes and a large SNR.

A list of the training data available for each phase of classification and the resulting feature vectors used for classification are provided in Table 8. Further details can be found in the demonstration report for FLBGR (Billings et al., 2007).

Table 8. Different phases of classification on the 20 mm RF and the RR and the corresponding training data and feature vectors selected for classification for the EM61 and EM63.

Phase	EM61		EM63	
	Training Data	Feature Vectors	Training Data	Feature Vectors
20mm RF: I (19-14)	Test plot	*Misfit/amplitude* Lsum	Test-plot, 10 emp. 37 mm in 19-14 Test-data 10 emp. 37 mm in 21-14	*Misfit/amplitude* ksum
20mm RF: II (21-14)	Test plot and 19-14	*Misfit/amplitude*^{2} Lsum	Test-plot, 20 emp. 37 and 19-14	*Misfit/amplitude*^{2} Ksum
RR: I (I-12, J-13)	Test-plot	*Misfit/amplitude* Lsum	Test-plot	*Misfit/amplitude* ksum
RR: II (L-14, K-15)	Test-plot and I-12, J-13	*Misfit/amplitude*^{2} Lsum	Test-plot and I-12, J-13	*Misfit/amplitude*^{2} log(k_1)
RR: III (I-13, J-12, L-13, L-14)	Test-plot, I-12, J-13	*Misfit/amplitude*^{2} Lsum	Test-plot, I-12, J-13	*Misfit/amplitude* log(k_1)

7.4.2 Classification Strategy at Camp Sibert

The classification strategy at Camp Sibert was constrained by ground truth information from the GPO and the results of excavating approximately 150 items on the site. For the magnetic a scatter plot (Figure 15a) showing remanence versus moment indicates that either parameter would provide a good basis for a discrimination strategy. However, the potential for significant remanence in the SEED items, led us to use the moment to prioritize digging order. Thresholds dictating which items to dig were selected by reference to the training data (Figure 15a).

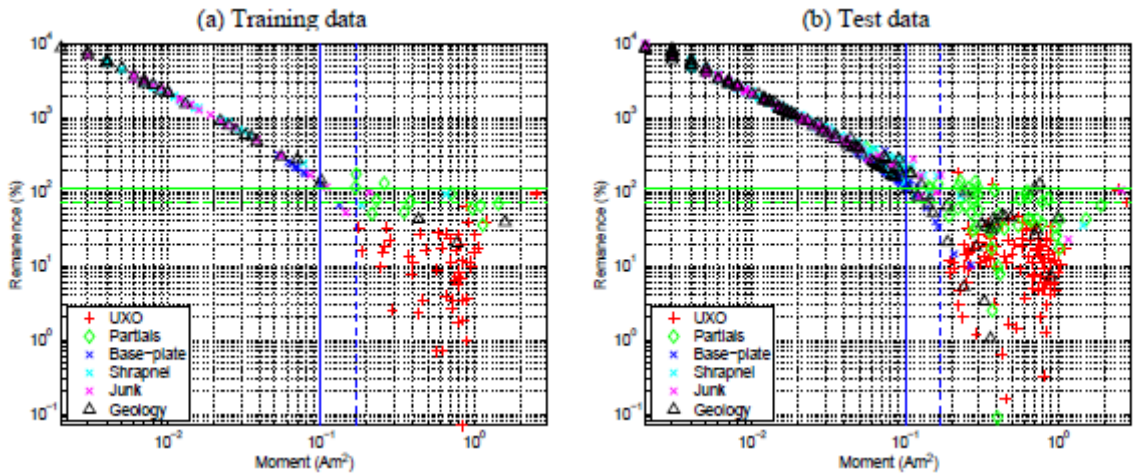

Figure 15. Scatterplot of the moment versus remenance for the (a) training data and (b) test data.

For the EM61 cart we trained a PNN classifier on all "passed" feature vectors from the GPO and ground truth datasets. We used L1(t1) and the maximum of L1(t1)/L3(t1) and L2(t3)/L2(t1) for the classification. The two feature vectors were first standardized so that they had zero mean and unit standardization. All non-UXO items were combined into a single class. The resulting classifier appears to be intuitively reasonable (Figure 16a). Small items with fast time-decays are highly unlikely to be UXO and can be safely left in the ground. At a given size value, the slower the time decay the more likely the item is a UXO.

The dig sheet will be ordered according to Pnon-UXO (the PNN probability the item belongs to the non-UXO class) with the item least likely to be a UXO appearing first. We then need to elect threshold values of Pnon-UXO for both the high and low FOM anomalies. For FOM = 1, we can be more aggressive in our selection of the threshold so we can stop digging high FOM anomalies sooner than low FOM anomalies. By investigation of the classification contours, we selected the cut-off values delineated in Figure 16a.

The classification strategies for the MTADS EM61 (Figure 16c) and MTADS EM61 cooperatively inverted (Figure 16e) were identical to that used for the EM61 (except different thresholds were chosen).

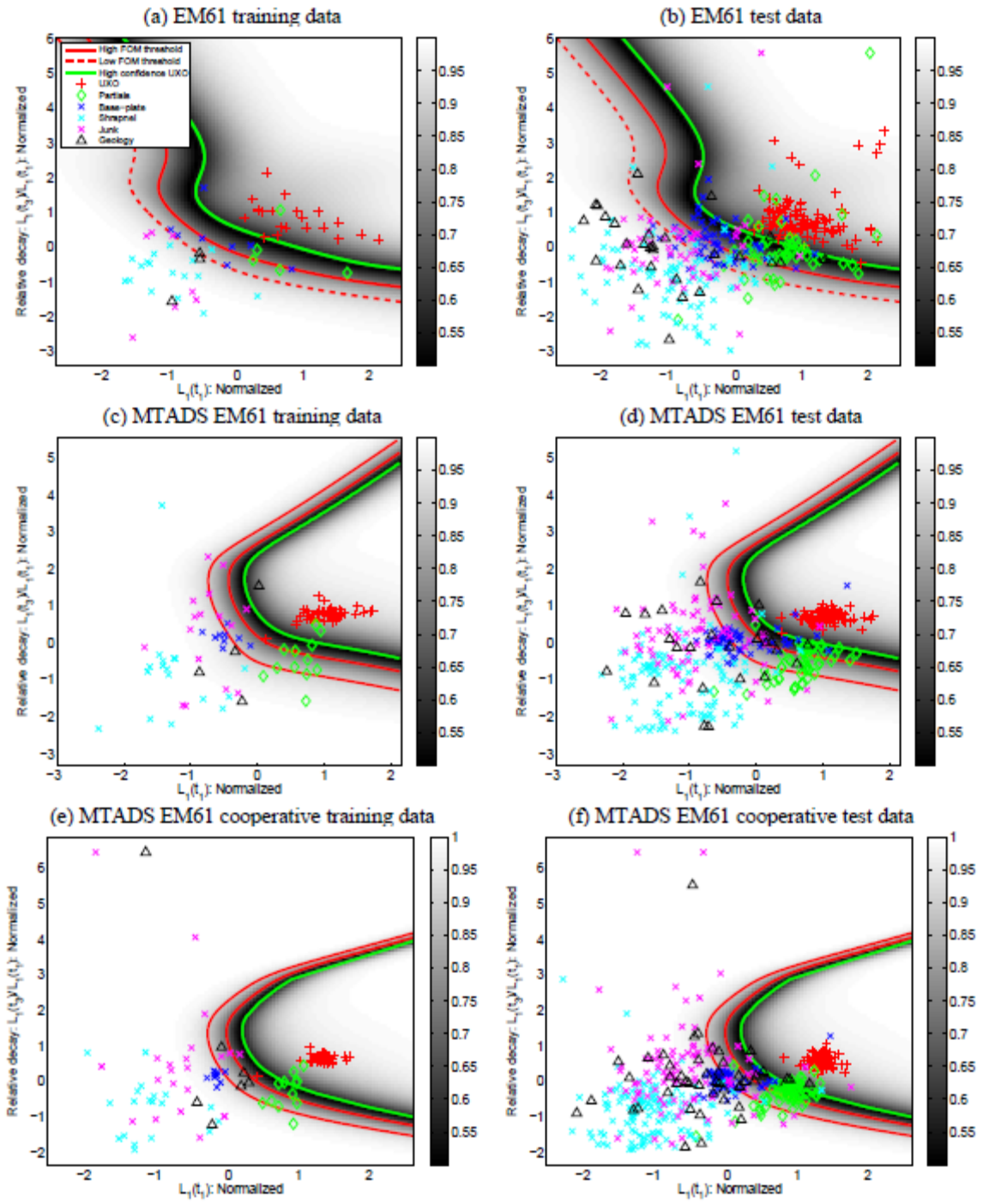

Figure 16. Feature vector plots for the contractor EM61 (a and b), MTADS EM61 (c and d), and MTADS EM61 cooperative (e and f) datasets.
The plots on left are for the training data, while those on the right are test data.

For the EM63, analysis of the GPO and the ground truth data indicates that a combination of a size-based and a decay-based parameter is sufficient to discriminate the 4.2-inch mortars from the nonhazardous items. Figure 17a shows a PNN classifier with a Gaussian kernel function that was trained on the log10(k1) and L1(t15)/L1(t1) features from the GPO and ground truth data. For the cooperatively inverted EM63, a PNN classifier was trained on the GPO and ground truth items using exactly the same parameters as the EM63 data inverted noncooperatively (Figure 17c). Qualitatively, the classification boundary looks very similar to the noncooperatively inverted data. For dig sheet creation, we used exactly the same thresholds on the PNN probabilities as was used for the noncooperatively inverted data.

Further details on the classification methods can be found in Billings et al. (2007).

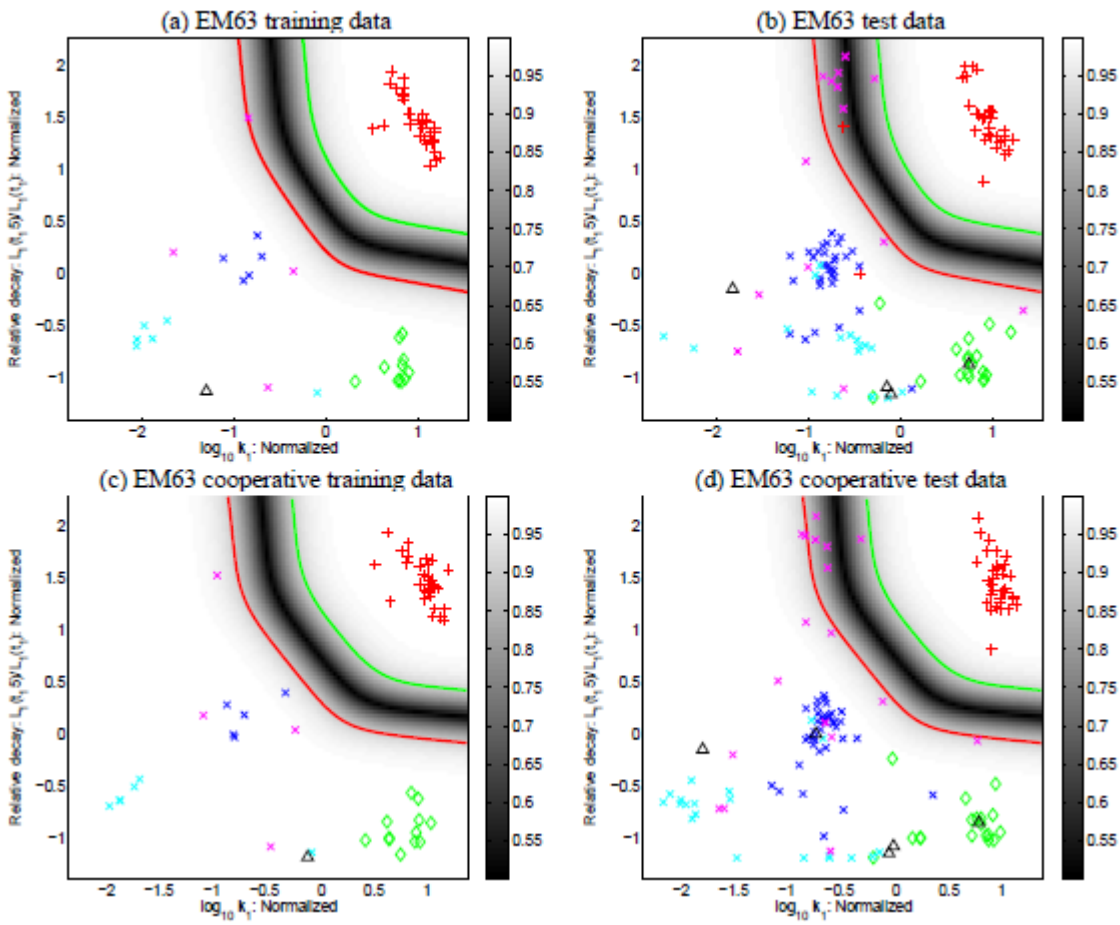

Figure 17. Feature vector plots for the contractor EM63 (a and b) and EM63 cooperative (c and d) datasets.
The plots on left are for the training data, while those on the right are test data.

7.4.3 Classification Strategy at Fort McClellan

For training data, we used the 21 test pit measurements (over seven different UXOs), the 18 items measured on the GPO and a random selection of 60 items from the 401 measured at the site.

The dig team listed an item as "Demo" if it contained or was suspected to contain energetic materials that made moving the round dangerous. Many of the 75 mm and 3.8-inch shrapnel rounds recovered from the site were missing the lead shot that had been blown out of the back of the round on detonation. While these rounds are no longer dangerous, they are indistinguishable (from a TEM perspective) from the rounds containing energetic materials. We therefore decided to call any 75 mm UXO scrap item of 7 lbs and any 3.8-inch UXO scrap item of 10 lbs an "item of concern."

From inspection of the feature space plots, we deduced that a combination of object size and time decay information would provide the most effective discrimination information (Figure 18). This is the same feature space used at Camp Sibert. After investigating an SVM classifier and a PNN, we eventually settled on a quadratic discriminant analysis classifier. Visual inspection of the decision surfaces indicated that it would provide the most effective discrimination strategy. By visual inspection of the classification surface, we settled on a decision surface where the target of interest (TOI) and non-TOI probabilities were equal. Everything inside the boundary (to the top right) is considered a TOI and would be excavated; everything outside the boundary (below and to the left) would be left in the ground. Further details on the classification methods can be found in Billings and Kingdon (2009).

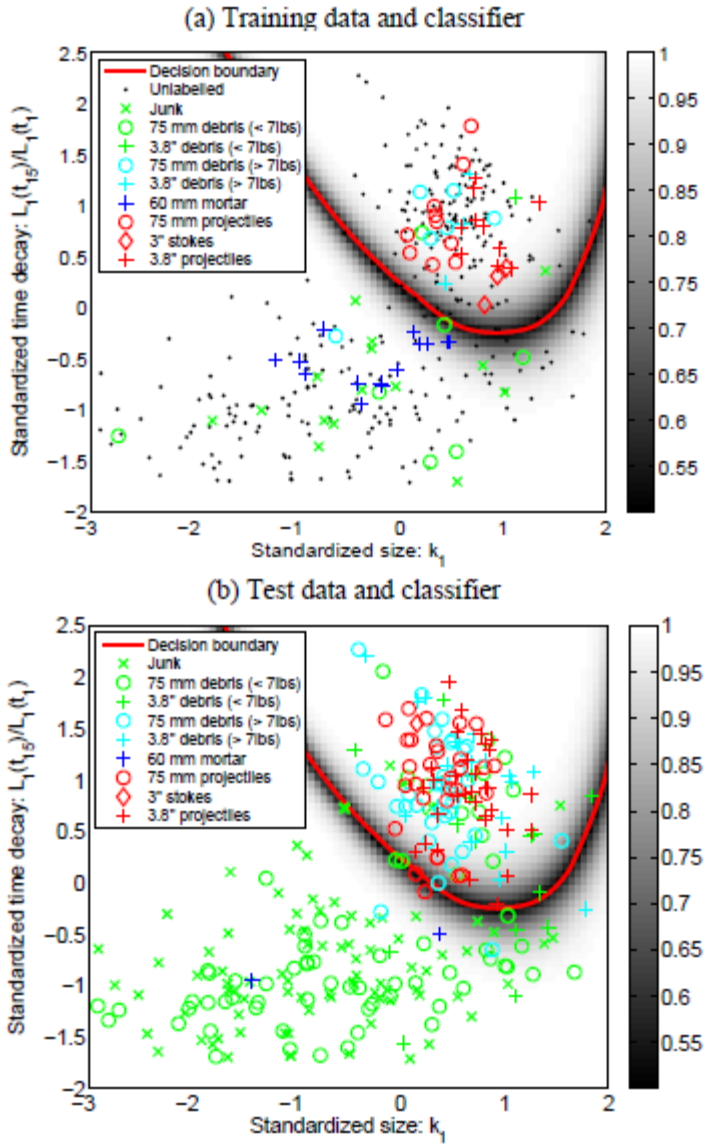

Figure 18. Quadratic discriminant analysis classification of training and test data.

7.5 DATA PRODUCTS

The most important final data product is a dig list, which consists of the locations of detected anomalies and a prioritized digging order, with a stop-dig point. Figure 19 shows the dig list format recommended by the Program Office that was used for the last two demonstrations.

- The first item in the list (Rank = 1) should be that which you are most certain does NOT need to be dug up (shown in green).
- The bottom items should be those that you are most certain are munitions and must be dug (shown in red). Thus, larger numerical rankings are associated with likely targets of interest.

- A threshold should be set at the point beyond which you would recommend digging all targets, either because you are certain they are ordnance or because a high confidence determination cannot be made (heavy black dividing line in Figure 19).
- Two other bands should be specified indicating (1) the range of targets where the SNR, data quality, or other factors prevent any meaningful analysis (shown in grey) and (2) the range of targets where the data can be fit in a meaningful way, but the derived parameters do not permit a conclusion (shown in yellow). These represent two levels of "guessing."

Rank	Comment
1	
2	High confidence NOT ordnance (no dig)
3	
…	Can't make a decision (dig)
…	
…	
…	Can't analyze (dig)
…	
…	Can't make a decision (dig)
…	
97	
.	High confidence ordnance (dig)
.	
N	

Figure 19. Ranked dig list.
High numbers represent likely UXO.

8.0 PERFORMANCE ASSESSMENT

Performance assessment at each of the three sites will now be discussed. We present the performance criteria and metrics for each site, then discuss the discrimination performance in greater detail. Table 9 lists the overall performance criteria and metrics used for evaluation.

Table 9. Performance criteria and the metrics used for evaluation at FLBGR.

Performance Criteria	Description	Primary or Secondary
Pd for EM63	(# of EM61 detected items detected with EM63) / (# of EM61 detected items)	*Primary*
Pdisc	(# of MEC items detected and recommended for excavation) / (# MEC items detected)	*Primary*
False alarm rate (FAR)	# of anomalies not corresponding to an ordnance item	*Primary*
Probability of false alarm (Pfa)	# FPs (i.e. declaration of ordnance) corresponding to clutter/# of opportunities for FP	*Primary*
Georeference position accuracy	Distance to interpreted items Primary	*Primary*
Terrain/vegetation restrictions	General qualitative observations on the suitability for the conditions encountered at the test site	*Primary*
Survey rate for magnetometer system	Hectares per day	*Secondary*
Survey rate for EM63 system	Hectares per day	*Secondary*
Percent site coverage	Percentage of area	*Secondary*
Processing time (for initial processing)	Total minutes of operator time per tile	*Secondary*
Processing time (interpretation)	Total minutes of operator time per anomaly	*Secondary*
Ease of use (hardware)	Number of people required to operate sensor and any support equipment. Skill level required by operators. Oversight required by site geophysicist. Problems detected by quality assurance (QA) officer	*Secondary*
Accuracy of inversion parameters	Comparison of spread in parameters for a given ordnance class for cooperative versus single inversion	*Secondary*

8.1 DISCRIMINATION PERFORMANCE AT FLBGR

In Table 10, we list the performance metrics and results achieved at FLBGR. In the sections that follow, we discuss the discrimination performance in more detail.

Table 10. Performance metrics at FLBGR.

Type of Performance Objective	Primary Performance Criteria	Expected Performance (Metric)	Actual Performance (Objective Met?)			
			EM61 RR	EM61 20mm RF	EM63 RR	EM63 20mm RF
Qualitative	Terrain/vegetation restrictions	Operator acceptance for use at the site	Yes	Yes	Met	Met
	Ease of use (hardware)	Operator and site geophysicist acceptance	Yes	Yes	Partly	Partly
Quantitative	Pd of EM63 sensor	≥ Pd for EM61 towed array	NA	NA	Unknown	Unknown
	Pdisc with a 50% reduction in false alarms	≥ 0.9	Yes	Yes	Yes	Yes
	FAR with Pdisc = 1	> 25% reduction in false alarms	Yes	Yes	Yes	Yes
	Location accuracy of interpreted items	<0.2m	Yes	No	Yes	Yes
	Survey rate for magnetometer system	1 hectare/day	Yes	Yes	Yes	Yes
	Survey rate for EM63 system	1/3 hectare/day	NA	NA	Yes	Yes
	Percent site coverage	>95%	Yes	Yes	Yes	Yes
	Processing time (initial processing)	< 1 day per tile (1 acre)	Yes	Yes	Yes	Yes
	Processing time (interpretation)	< 5 minutes operator time per anomaly	Unknown	Unknown	Unknown	Unknown
	Accuracy of inversion parameters	Within class variance of cooperative < single inversion	NA	NA	NA	NA

Table 11 summarizes the discrimination performance for each phase of the demonstration. We exclude the Phase II results at the RR as there was only one UXO item found. In addition, we only include anomalies where the inversion returned an acceptable feature vector and list the "failed fits" in a separate column. The table lists the false alarms, the number of UXO recovered and the Pd for the selected operating points of the amplitude and classification type methods. Note that we express our discrimination results using Pd and not Pdisc. The difference is that Pdisc considers performance only on anomalies above the selected detection threshold, whereas Pd considers all anomalies for which ground truth is available (and there is a valid fit to the data). This was done to emphasize that the classification method, at times, ranked low amplitude anomalies quite high in the priority list. We can draw the following conclusions regarding performance from the table:

- At the 20 mm RF, we tended to choose an operating point that was too aggressive and consequently would have missed a number of UXO. However, in each case these extra UXO were included not too much further down the dig list, and all

UXO were recovered with many fewer excavations than the amplitude-based method.

- At the 20 mm RF, the operating point of the amplitude-based method was also too aggressive, and a number of UXO would have been missed.
- For each phase, we meet the objective that Pdisc>0.9 at the point where we dig, 50% fewer false alarms than the amplitude-based method. In fact, we achieve Pdisc=1 for each phase.
- For each phase, we meet the objective of at least a 25% reduction in false alarms (compared to the amplitude method) at the point where Pdisc = 1. The smallest reduction is 64% with the Phase III RR results achieving a 90% reduction in false alarms.

Table 11. EM61 discrimination performance results at the 20 mm RF and RR.

Phase	# UXO (with valid fits)	Amplitude			Classifier						
		False alarms (OP)	UXO (OP)	Pd (OP)	False alarms (OP)	UXO (OP)	Pd (OP)	Pd 50% false alarms	False alarms Pd=1	False alarm reduction Pd=1	Failed fits
Phase I 20mm RF	8	42	6	75%	6	6	75%	1	8	81%	28 (0)
Retrospective	8	42	6	75%	8	8	100%	1	8	81%	28 (0)
Phase II 20mm RF	9	59	8	89%	13	6	75%	1	21	64%	31 (1)
Phase I RR	6	84	6	150%	26	6	100%	1	26	69%	13 (2)
Retrospective	6	84	6	150%	24	6	100%	1	24	71%	13 (2)
Phase III RR	20	126	20	100%	36	20	100%	1	13	90%	68 (2)

OP = operating point

For Phase I at both sites we show the actual (using *misfit/amplitude*) and retrospective performances (using *misfit/amplitude*^{2}). The false alarms, number of UXO, and Pd at the operating points of the amplitude and classification methods are shown. The failed fits column lists the number of inversions with failed fits with the number of UXO shown in brackets. Pd at the point with a 50% reduction in false alarms and the reduction in false alarms at Pd = 1 are also shown.

Table 12 lists the performance results for the EM63 at the 20 mm RF and the Rocket Rage. This performance metric is met for all phases. The results for Grid 21-14 include two 37 mm projectiles at 40 cm that were below the production threshold. The Pd of the discrimination method includes these two projectiles and indicates that 110 false alarms are required to achieve Pd=1. This is the same number as the number of false-alarms with the production method. The difference is that the production method only recovers six UXO at that threshold compared to eight for classification. To recover those same six UXO, the classification method requires only 36 false alarms (thus Pdisc = 1 at the point where there is a 50% reduction in false alarms).

Table 12. EM63 discrimination performance results at both sites.

Phase	# UXO (valid fits)	Amplitude			Classifier						
		False alarms (OP)	UXO (OP)	Pd (OP)	False alarms (OP)	UXO (OP)	Pd (OP)	Pd 50% false alarms	False alarms Pd=1	False alarm reduction Pd=1	Failed fits
Phase I 20 mm RF	7	107	7	100%	37	7	100%	1	11	90%	28 (1)
Phase II 20 mm RF	8	110	6	75%	37	6	75%	0.75	110	0%	55 (2)
Phase I RR	6	84	6	100%	26	6	100%	1	26	69%	37 (2)
Retrospective	6	84	6	100%	24	6	100%	1	24	71%	37 (2)
Phase III RR	21	143	21	100%	28	21	100%	1	22	85%	77 (1)

For Phase I at the RR we show the actual (using *misfit/amplitude*) and retrospective performances (using *misfit/amplitude*^{2}). For a description of the columns, refer to Table 10.

8.2 PERFORMANCE METRICS AT CAMP SIBERT

In Table 13 we list the performance metrics and results achieved at Camp Sibert. In the sections that follow, we discuss the discrimination performance in more detail. For magnetics, the training data revealed a number of 4.2-inch mortars with relatively large remanence (Figure 15a). All 4.2-inch mortars had moments greater than 0.17 ampere-meter squared (Am^2), and could therefore be considered "large." Consequently, the size of the moment (and not the remanence) was used to prioritize the dig list. We also produced a remanence prioritized dig list so we could test performance using that discrimination metric.

All recovered moments from 4.2-inch mortars in the blind test data were above the stop-digging threshold (or were listed as can't-analyze), whereas two items in the remanence ranked list lay outside the region recommended for excavation (Figure 15b). There were 95 items in the blind test data classified as can't-analyze. The receiver operating characteristic (ROC) curve for the dig list ranked by moment is shown in Figure 20a (including can't-analyze) and Figure 20b (excluding can't-analyze).

Table 13. Expected performance and performance confirmation methods.

The georeference metric consists of the percentage of items within 20 cm of the ground truth location. Metrics that were successfully met are shown in green, while those that were not are shown in yellow.

Performance Criteria	Expected Performance Metric	Performance Confirmation Method	Magnetometer Moment	Magnetometer Remanent	EM61 Size	EM61 Classifier	MTADS EM61 Size	MTADS EM61 Classifier	MTADS EM61 Cooperative	EM63	EM63 Cooperative
Pdisc (recovered) at operating point	≥0.95	Compare to ground truth	1.00	0.98	1.00	1.00	1.00	1.00	1.00	1.00	1.00
False-alarm rate at operating point	Not specified	Compare to ground truth	0.28	0.24	0.67	0.62	0.58	0.56	0.26	0.33	0.18
False-alarm rate with PDisc = 1	FA reduced >25%	Compare to ground truth	0.24	0.31	0.48	0.41	0.55	0.48	0.28	0.28	0.14
Geo-reference position accuracy	<0.2m	Compare to ground truth	85%	86%	61%	61%	66%	66%	82%	86%	89%
Processing time (interpretation)	< 5 minutes per anomaly	Data analysis log	3.7	3.7	9.1	9.1	5.0	5.0	7.4	5.3	8.8
Inversion accuracy	Cooperative better than single	Compare to ground truth	NA	NA	NA	NA	NA	NA	Yes for size	NA	Yes

Table 14. Summary of the discrimination performance of the nine different sensor combinations or discrimination methods. Numbers in parentheses represent the results excluding the can't-analyze category, while the numbers in parentheses in the can't-analyze column show the number of UXO in the can't-analyze category.

Technologies	# Alarms	# UXO	At Operating Point				At Pdisc = 1	
			# Can't-Analyze	# FPs	Pdisc	Pfa	# False Alarms	Pfa
Magnetometer (moment)	825	118	95 (2)	198 (103)	1	0.28	170 (75)	0.24
Magnetometer (remanence)	825	118	95 (2)	170 (75)	0.98	0.24	218 (123)	0.31
EM61 (size)	546	118	149 (13)	285 (136)	1	0.67	205 (56)	0.48
EM61 (classification)	546	118	149 (13)	264 (115)	1	0.62	174 (25)	0.41
MTADS EM61 (size)	734	119	285 (8)	357 (72)	1	0.58	338 (53)	0.55
MTADS EM61 (classification)	734	119	285 (8)	344 (59)	1	0.56	293 (8)	0.48
MTADS EM61 cooperative	734	119	205 (2)	275 (70)	1	0.45	208 (3)	0.34
MTADS EM61 cooperative (with magnetics)	983	119	226 (2)	307 (81)	1	0.26	240 (14)	0.28
EM63	150	34	30 (0)	38 (8)	0.97	0.33	47 (17)	0.41
EM63 cooperative	150	34	16 (0)	21 (5)	1	0.18	16 (0)	0.14

8.2.1 Discrimination Performance of Full-Coverage Data Sets

In this section we investigate the performance of the following three methodologies:

- Statistical classification of the EM61 data
- Statistical classification of the MTADS EM61 data
- Statistical classification of the MTADS EM61 cooperatively inverted data.

For each of these sensor combinations we used a size-based and a time-decay-based feature vector and a PNN classifier. Figure 16 plots the two feature vectors for each sensor combination over the decision surface and has separate plots for the training and test data. Feature vectors within the UXO, partial rounds, and base-plate classes are more tightly clustered for the MTADS EM61 than for the contractor EM61, with further improvement evident in the MTADS EM61 cooperatively inverted. For all three data sets we used two decision boundaries: a more aggressive one for high FOM anomalies and a less aggressive one for low FOM anomalies. For each of the three data sets, all 4.2-inch mortars lie on the appropriate side of the "high confidence UXO" decision boundary. As shown in Figure 20, all potentially hazardous items are recovered with very few FPs (ignoring the can't-analyze category) with FP = 25, 8, and 3 for EM61, MTADS EM61, and MTADS EM61 cooperative, respectively. The discrimination performance of both MTADS EM61 data sets are degraded by the high number of can't-analyze items: 285 in the MTADS EM61, compared to 205 in the MTADS EM61 cooperative. The lower number of can't-analyze items in the cooperatively inverted data occurs because the magnetics data removes some of the depth ambiguity inherent in the EM data.

Figure 20 compares the magnetometer, EM61, MTADS EM61, and MTADS EM61 cooperative ROC curves. When including the can't-analyze category (Figure 20a), the magnetometer data requires the least number of FP excavations at its operating point (170 compared to 264, 344 and 275 for the EM61, MTADS EM61 and MTADS EM61 cooperative). 72% of the non-UXO items can be left in the ground with the magnetometer data compared to 33, 44 and 55% for the other datasets. When excluding the can't-analyze category (Figure 20b), it is evident that feature vectors extracted from the MTADS EM61 and MTADS EM61 cooperatively inverted data are more highly discriminatory than the magnetometer or EM61.

The Institute for Defense Analyses (IDA) demonstration protocol involved scoring the cooperatively inverted data on the union of the EM and magnetometer data sets. There were 249 magnetometer anomalies that did not have a corresponding EM anomaly. Only 32 of these required excavation using the same criterion utilized in the moment ranked dig sheet, with 21 of those anomalies in the can't-analyze category. This modest increase in the number of anomalies to excavate using cooperative inversion is more than compensated by the 80 fewer can't-analyze anomalies.

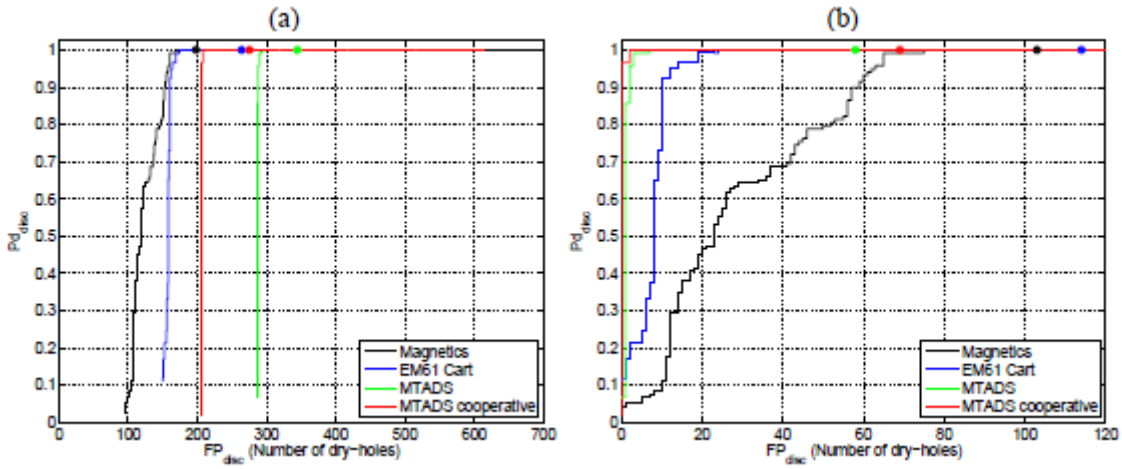

Figure 20. Comparison of ROC curves for the magnetics (moment), contractor EM61, MTADS EM61, and MTADS EM61 cooperative data sets: (a) including can't-analyze category and (b) excluding can't-analyze anomalies.

8.2.2 Discrimination Performance of Cued-Interrogation Data Sets

In this section we investigate the performance of the following two cued-interrogation methodologies:

- Statistical classification of the EM63 data
- Statistical classification of the EM63 cooperatively inverted data.

As with the EM61-based data sets, discrimination was based on a PNN classifier with a size-based and a time-based feature vector. The size-based feature vector was the k1 parameter from the primary polarization with the ratio of primary polarizations at the 15th and 1st time channels used as the time-decay parameter. The UXO, partial rounds, and base-plate classes are tightly clustered in both data sets, with less variation in the cued-interrogation data (Figure 17). The UXO class for the EM63 data contain two outliers, one of which causes the false negative evident in Figure 21a and c. The EM63 when cooperatively inverted produces a "perfect" ROC curve with 0 FPs (excluding can't-analyze anomalies) at the point where all UXO are recovered. At the operating point, a total of 21 FP, with 16 of these in the can't-analyze category, are required, with 82% of non-UXO left in the ground. In Figure 14 we also show an ROC curve for the MTADS EM61 cooperative when restricted to the same 150 cued-interrogation anomalies. It also results in a "perfect" ROC curve but requires 34 FPs at the operating point, with 6 of those in the can't-analyze category.

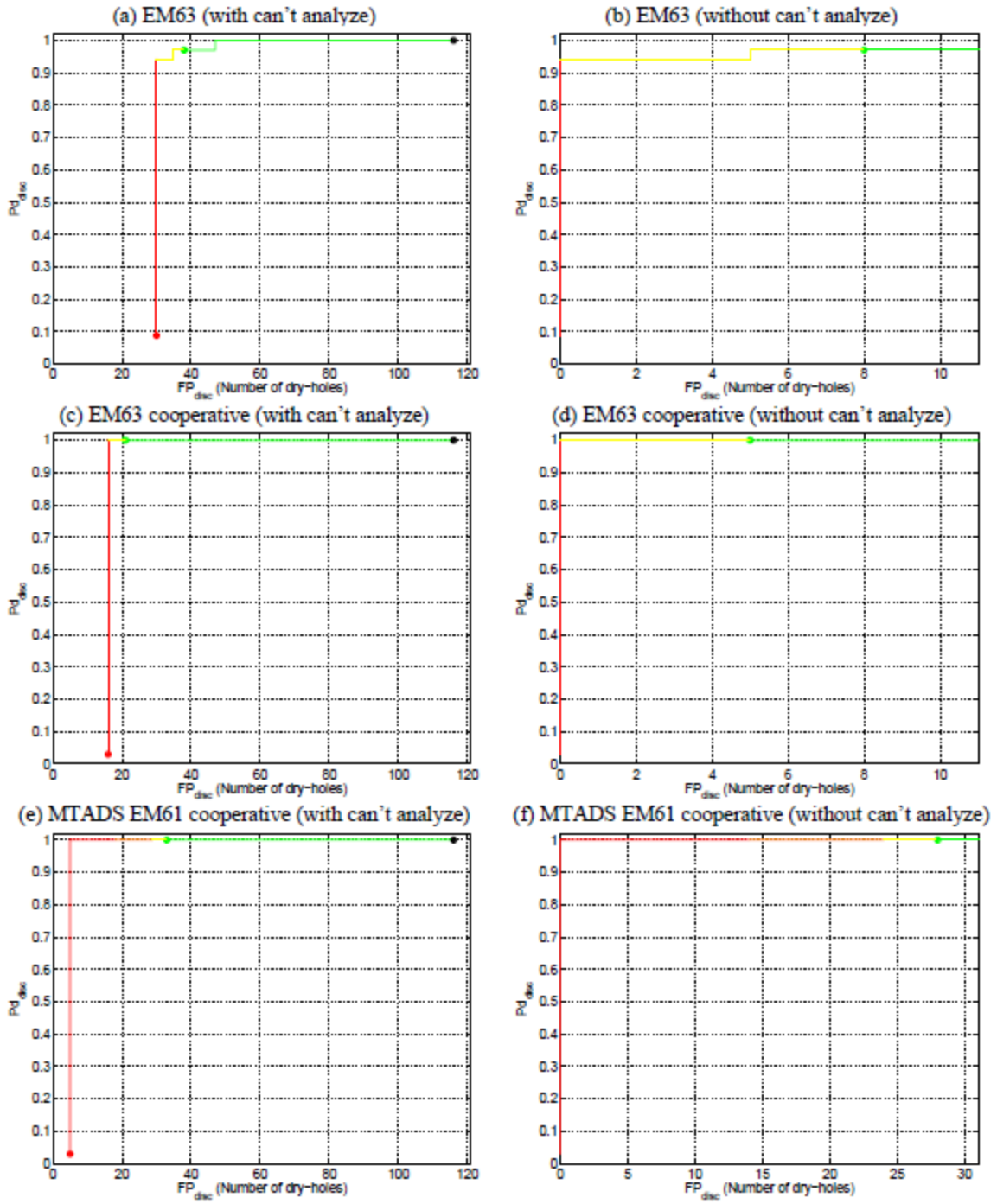

Figure 21. ROC curves for the EM63 (a and b), EM63 cooperative (c and d), and MTADS EM61 cooperative (e and f) on each of the cued-interrogation anomalies. The plots on left include the can't-analyze category, while those on the right exclude them.

8.2.3 Processing Time Required

During the demonstration, each analyst kept a log of the time spent on each step in the inversion process (Table 15). These steps included:

- *Preprocessing and setup.* This included all the time spent manipulating and importing into UXOLab all the data provided by each data collection demonstrator. Times ranged from 4 hours for the EM63 to 40 hours for the contractor EM61. A significant amount of time had to be spent on the EM61 cart data as the delivered data were not as well organized as the MTADS EM61 and magnetometer data. For the cooperative inversion, we assumed the times would be the same as for single-inversion.
- *Defining masks.* This included the time spent automatically generating and then reviewing each mask prior to inversion. Typically this required 6 to 8.5 hours of analyst time for the full-coverage datasets. We assumed mask times for cooperative and single inversion were identical.
- *Inversion.* This is computer time required to invert all anomalies, which ranged from 3 hours for the magnetics to 15 hours for the MTADS EM61. The cooperatively inverted datasets required less computer time as many fewer start-models were required.
- *Portable Document Format (PDF) report generation.* This is computer time required to create PDF reports for each dataset. These are rough estimates as many of these were left to run overnight.
- *Quality control (QC) of fits using the PDF report.* The analyst reviewed each anomaly and determined if the inversion result could be trusted. Times ranged from 7 hours for the MTADS EM61 cooperative to 26 hours for the EM61. The better the data quality, the less time required for QC.
- *Mask and invert second pass.* This step required 6 to 8.5 hours for the full-coverage data sets. No remasking was required for the EM63 cooperative.
- *QC using second PDF report.* This step ranged from 4 to 10.5 hours for the full-coverage data sets.
- *Independent QC of fits.* The QC operator spent between 8 and 10 hours on QC of the fitted data. This included any time required to remask and reinvert anomalies that failed the QC check.
- *Discrimination.* The selection of feature vectors, training of the classifier, application of the classifier, and creation of the dig sheet typically required about 6 hours per data set.

Table 15. Time spent processing each of the different sensor combinations.

Operation	Magnetmeter	EM61	MTADS EM61	MTADS EM61 Cooperative	EM63	EM63 Cooperative
Preprocessing and setup	11	40	18.25	18.25	4	4
Defining masks	7	6.5	8.25	8.25	1.25	1.25
Inversion	3	6.5	15	4	7	2
Generating PDF report	20	6.5	18.5	18.5	2.25	2.25
QC of fits using PDF report	18.75	26	15.75	7	4	6
Mask & invert 2nd pass	7.5	8.5	6.5	6	0.5	
Create 2nd PDF report	5	5	3	4	1	
QC using 2nd PDF report	8	4.25	10.5	5	0.5	
Independent QC of fits	8.25	10	10	10	6	6
Discrimination	2	6	6	2	3	1
Total (hours)	90.5	119.3	111.8	83.0	29.5	22.5
Total minus computer time (hours)	62.5	101.3	75.3	56.5	19.3	18.3
Number of targets	1007	671	908	908	216	216
Per target (min)	5.4	10.7	7.4	5.5	8.2	6.3
Per target minus computer time (min)	3.7	9.1	5.0	3.7	5.3	5.1
Per target minus computer & setup (min)	3.1	5.5	3.8	2.5	4.2	4.0

We had expected to be able to process and interpret each anomaly in less than 5 minutes. As shown in Table 15, if we include the set-up time, then this time goal was met with the magnetometer and MTADS EM61 data but not with any of the other data sets. The contractor EM61 required the most time with 9.1 minutes per anomaly. The MTADS EM61 and EM63 cooperative inversions both required about 8.75 minutes per anomaly (the times in Table 11 did not include the time required to invert the magnetometer data). If we discount the time required for setup (not unreasonable if the data are delivered in a more usable form), then the time goal is met by all the single inversion methods except the EM61. The cooperative inversions then require about 7 minutes per anomaly.

8.3 PERFORMANCE METRICS AT FORT MCCLELLAN

Performance metrics and performance confirmation methods used at Fort McClellan are listed in Table 16. Additional details on the discrimination performance are provided in the following discussion.

Table 16. Expected performance and performance confirmation methods.

Performance Criteria	Expected Performance Metric (pre-demo)	Performance Confirmation Method	Actual (post-demo)
Quantitative Criteria			
PDisc (recovered) at operating point	> 0.95	By reference to validation information and ranked dig sheet	0.96
False alarm rate with PDisc = 0.95	>50% reduction in false alarms	By reference to validation information and ranked dig sheet	86%
False alarm rate with PDisc = 1	> 25% reduction in false alarms	By reference to validation information and ranked dig sheet	66%
Within class variation of $\log_{10}L_1(t_1)$	< 1	By reference to validation information and inversion parameters	0.96 & 1.12 (for 75mm and 3.8" TOI)
Within class variation of $L_1(t_{20})/ L_1(t_1)$	< 25%	By reference to validation information and inversion parameters	83 & 80% (for 75mm and 3.8" TOI)
Estimated depth	90% within < 0.15 m	Comparison of estimated depth with ground truth depth	60% TOI within 0.15 m
Processing time (interpretation)	Less than 10 minutes per anomaly	Entries in data analysis log	Poor records kept
Survey rate	> 30 items/day	Entries in data collection log	25.5
Qualitative Criteria			
Reliability and robustness	Operator acceptance	General observations	No

8.3.1 Discrimination Performance at Fort McClellan

The dig sheet was created using the rules described above using the probability of TOI to create the ordering. The first item in the list has the smallest probability of being a TOI and hence is the item least like to be ordnance. In the dig list submitted to the Program Office, there were 33 can't-analyze items, 147 high confidence TOI and 161 high confidence NOT TOI. Figure 22 (a) and (b) show two different ROC curves generated once the ground truth was released. The first ROC curve Figure 22(a) was generated assuming that only 10 lb 75 mm and 15 lb 3.8-inch shrapnel rounds were TOI. The second ROC curve Figure 22(b) was generated assuming any 75 mm UXO scrap of 7 lb or greater and any 3.8-inch UXO scrap of 10 lbs or greater was a TOI. We left in the two QC fails (items 80 and 392) as well as the item we suspect of having incorrect ground truth (item 5). Comparing the two ROC curves, we find (first ROC curve number is listed first):

- 64 compared to 119 TOI in the test-data
- 245 compared to 190 non-TOI in the test-data
- 62 (97%) TOI compared to 114 (96%) TOI recovered at the operating point
- 86 non-TOI compared to 34 non-TOI excavated at the operating point

- 93 non-TOI compared to 64 non-TOI (51 non-TOI if we were to exclude item 5 from consideration) excavated at the point where all TOI are recovered.

Apart from setting the operating point too early, the discrimination ranking is quite efficient with 63% (first ROC) and 66% (second ROC) of non-TOI left in the ground at the point where all TOI are recovered. If we exclude item 5, then 73% of non-TOI are left in the ground.

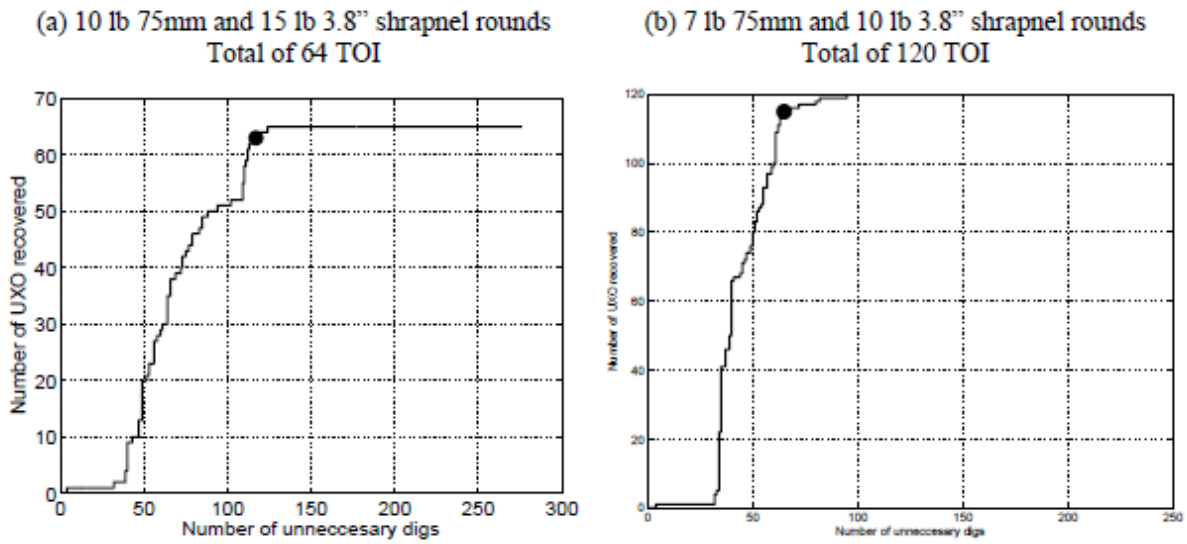

Figure 22. ROC curves corresponding to the quadratic discriminant analysis classifier of Figure 18.

The ROC curve in (a) was generated assuming only 10 lb 75 mm and 15 lb 3.8-inch shrapnel rounds were TOI. The ROC curve in (b) was generated assuming any 75 mm MEC scrap of 7 lb or greater and any 3.8-inch MEC scrap of 10 lb or greater was a TOI.

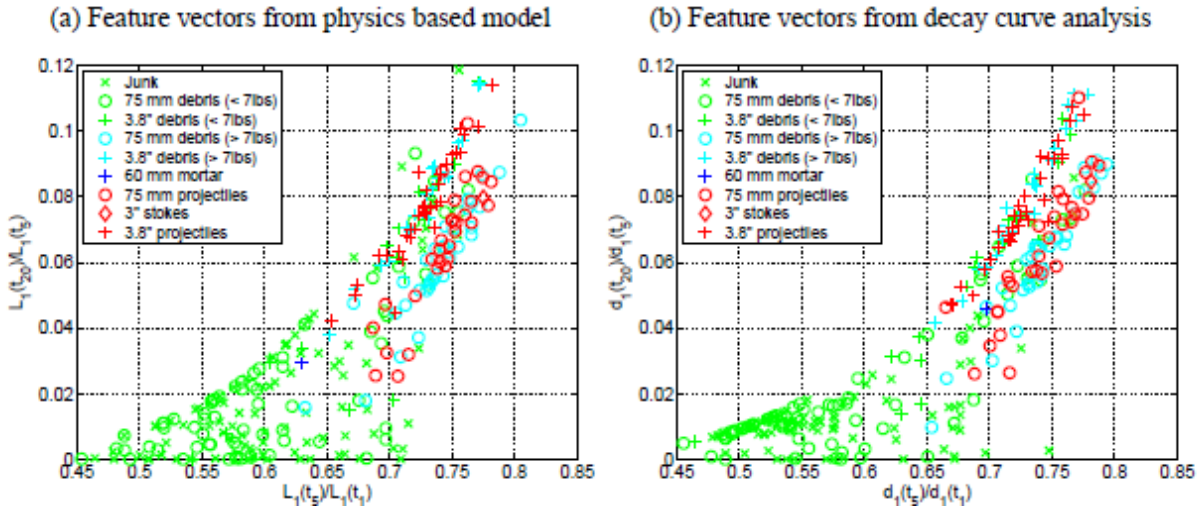

Figure 23. Comparison of time decay parameter feature spaces obtained from (a) inversion of a physics based model and (b) from decay curve analysis.

8.3.2 Number of Anomalies Covered Per Day

Figure 24 plots the number of anomalies that were surveyed each day. Apart from the first week, when instrument difficulties limited productivity and days when surveying was halted for demo shots, we always surveyed at least 20 items per day. Neglecting the days of incomplete or unattempted surveying, the average survey rate was 25.5 items per day.

Accurate records of the processing time per anomaly were not kept, but they should be comparable to the times recorded for the Camp Sibert EM63 analysis.

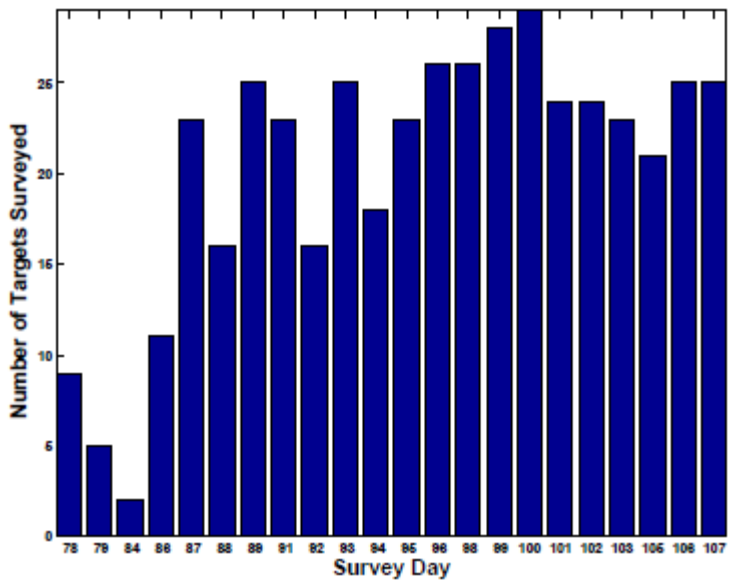

Figure 24. Number of cued interrogation anomalies surveyed each day.

9.0 COST ASSESSMENT

9.1 COST SUMMARY

As described previously, the data collection activities performed over the duration of this project include standard EM technologies (EM61 and EM63) and a cued-interrogation EM63 system. These systems were deployed to collect data in support of development and testing of discrimination techniques. As these activities have been performed at different sites and reflect continually advancing data collection and discrimination techniques, we present a summary and analysis of the costs for individual demonstrations as well as comparisons of the costs where applicable.

The Camp Sibert demonstration consisted of analysis of data collected by all participants in the studies. These tasks also included preparatory activities to update and modify UXOLab code to analyze and document the discrimination techniques more efficiently and to prepare to import data from new sensor technologies.

The total costs incurred for the data collection and discrimination activities performed for the various demonstrations are summarized in Table 17. Detailed breakdown of costs associated with each task are provided in the remainder of this section.

Table 17. Cost summary for all project activities and demonstrations.

Demonstration	Activities	Comments	Total Costs
Project initiation, software functionality	Setup, software development, and testing		$241K
FLBGR	EM63 data collection, EM61 and EM 63 data processing, analysis		$243K
Camp Sibert	EM 63 cued-interrogation data collection, processing, analysis		$99K
	Discrimination and analysis—all technologies		$248K
Fort McClellan	EM 63 cued-interrogation data collection, processing, and analysis		$190K
Other project activities	Misc. Program Office support activities, IPRs,* presentations, general management activities		$382K

The cost model for data collection tasks performed at FLBGR, Camp Sibert, and Fort McClellan are summarized in Table 18. These costs reflect those incurred to perform each demonstration; however, because there were some shared costs between other ongoing Sky projects and the discrimination studies directed by the Program Office and a varied scope of activities, these costs do not necessarily represent a cost or level of effort that should be assumed to be representative of future production-level efforts.

Table 18. Cost model for data collection demonstrations.

Cost Element	Notes, Data Tracked During Demonstration	Costs ($K)
Instrument costs	EM63 cued-interrogation system and magnetometer array were developed under ERDC and other projects. The production geophysics EM61 array used at FLBGR required no modifications. There were no incurred instrument development costs for this project. **FLGBR:** EM63 borrowed from CSM, rental required for part of demonstration. EM61 data previously collected under USACE–ERDC project. **Camp Sibert:** Borrowed EM63 **Fort McClellan:** Borrowed EM63	No project costs incurred.
Pre-demonstration testing	**FLBGR:** Costs for EM61 validation/testing of detection and discrimination characteristics, including emplacement of seed items for validation. The actual data collection was performed as part of ERDC-USACE projects. **Camp Sibert:** Pre-demonstration testing performed in Ashland, OR. **Fort McClellan:** Pre-demonstration testing performed in Ashland, OR.	$45.5 $21.2 $20.5
Mobilization and demobilization	**All Demonstrations:** Travel 1 day each way for field team, 1 day site setup and 1 day equipment packing & shipping **FLGBR:** Field team mobilization/demobilization from Vancouver to Denver **Camp Sibert:** 2 staff drove equipment from OR to AL; PI travel from Vancouver to AL. **Fort McClellan:** 1 person mobilized from Vancouver to AL.	$47.4 $17.0 $14.0
Site preparation	Site preparation performed under other projects or through the Program Office for Discrimination Studies.	No project costs incurred.
Survey costs	**FLBGR:** Labor, per diem, subcontractor support, Sky magnetometers, DAS and GPS equipment rental costs incurred. **Camp Sibert:** Labor, per diem, Sky DAS, and GPS equipment rental costs incurred, plus field supplies. **Fort McClellan:** Labor, per diem costs for 1 Sky geophysicist. NAEVA geotechnician support was provided, so no project costs incurred for part of the field team.	$62.5 $35.9 $41.4
Detection data processing	**FLBGR:** EM 63 and magnetics processing only; EM61 processing conducted as part of pre-demonstration testing. **Camp Sibert:** Processing was not tracked separately, estimated costs assuming 40% of total processing and discrimination. **Fort McClellan:** Includes senior geophysicist code modification, development, and QC	$36.1 $6K $31.0
Discrimination data processing	**FLBGR:** Includes discrimination and QA/QC **Camp Sibert:** Estimated as 60% of processing and discrimination **Fort McClellan:** Discrimination and analysis	$15K $8.2K $4.4K
Report preparation & management	**FLBGR:** **Camp Sibert:** **Fort McClellan:**	$27.8K $6.9K $19.2K

9.1.1 FLBGR Cost Summary

The costs incurred to perform the FLBGR discrimination demonstration are detailed in Table 19 and include:

- Pre-demonstration validation and testing of the EM61 discrimination capabilities using data collected at the RR
- Project team and equipment mobilization, data collection, and demobilization for magnetometer and EM63 data (included subcontractor support for magnetometer data collection)
- Data processing and QC
- Discrimination and data interpretation
- Ground truth excavation, including subcontractor EOD tech support
- Demonstration report preparation.

The costs associated with the FLBGR demonstration were shared with ongoing ERDC-USACE projects, so to assess the point at which discrimination becomes cost effective relative to the current practice of excavating each anomaly identified in a production EM survey, we performed an analysis using projected costs for both the current practice and utilizing discrimination to reduce the number of anomalies that must be excavated. The additional costs of data collection, processing, and interpretation, need to be offset by a reduction in the number of holes excavated. In this section of the report, we make a comparison of the expected costs of discrimination-based methods (EM61 and EM63) against a standard production survey with the EM61. The costs for a standard detection mode survey include:

- Site preparation (e.g., vegetation clearance, surface sweep)
- Geophysical survey
- Data processing and interpretation
- QA/QC
- Anomaly relocation (as appropriate)
- Anomaly excavation and documentation.

The costs for the discrimination strategy include the above, minus the costs for anomaly relocation and excavation for those anomalies not slated for excavation, plus the costs for:

- Additional geophysical surveying (as required);
- Additional data processing and interpretation; and
- Additional documentation and QC.

Table 19. FLBGR demonstration survey costs.

Categories	Pre-Mobilization Testing and Validation	Mobilization, Site Setup, & Demobilization	Data Collection	Data Processing	Data Interpretation & QC	Digging	Reporting	Total
Labor and subcontractors	$22,325	$23,680	$27,462	$33,887	$25,014	$15,331	$26,958	$174,657
Equipment	$ 0	$3204	$18,146	$ -		$ -	$	$21,350
Travel	$1167	$2384	$13,903	$ -		$665	$	$18,119
Materials and supplies, shipping	$ 0	$3909	$382	$ -		$ -	$	$4291
Std ODCs	$213	$1656	$2656	$4226	$682	$186	$825	$10,444
Total	**$23,705**	**$34,833**	**$62,549**	**$38,113**	**$25,696**	**$16.182**	**$27,783**	**$228,861**

For this demonstration we were able to accurately track the time (and hence costs) required for geophysical survey and initial processing. For the feature extraction and classification, we made some adjustments to the underlying inversion methods and trialed a number of different inversion strategies (noise-levels, masks, etc.) before settling on our final approach. Therefore, for these tasks we can only make estimates of reasonable times required for processing and interpretation.

For the cost comparison we estimate the fully burdened costs to the government for the geophysical and excavation phases of the work. We make the following assumptions:

- Mobilization/demobilization
  - The survey requires a 2000 km mobilization from Denver.
  - Mobilization includes 1 day preparation, 1 day setup on site and 1 day for a test plot survey (and associated processing).
  - Demobilization includes 1 day of packing up on site and 1 day of organization back at home base.
- The survey area is 100 acres.
- There are 100 anomalies per acre.
- Anomaly reacquisition proceeds at 150 items per day and requires a three-person crew (two field technicians, one UXO technician).
- It costs $100 to excavate each anomaly.
- EOD escort rate and per diem are the same as the field-technicians.
- There is a management cost of 10% of for the geophysical portion of the work (the excavation portion is embedded within the cost per anomaly).
- Daily rates are fully burdened and assume a 10-hour working day.
- Processing requires a Geophysicist II level analyst while interpretation is 50% at the same level and 50% at the next highest level.
- Per diem and hotel are assumed to total $150 per person per day, with 1.4 days of per diem/hotel per day in the field (5 day week, 2 days on the weekend).
- Equipment charge includes all equipment, vehicles (assumed 1 rental vehicle required) consumables, and supplies.
- Anomaly reacquisition costs are calculated assuming two field-technicians and an EOD tech can reacquire 150 anomalies per-day.
- Additional assumptions on each survey method are provided in Table 20.

Table 20. Assumptions used to compare the different survey methods.
The excavation reduction numbers were estimated from the performance at the RR.

Attribute	EM61 Production	EM61 Discrimination	EM63 Discrimination
Survey rate	10 acres per day	8 acres per day	1 per day
Equipment	5 EM61, 1 GPS, tow-vehicle/sled, 2 vehicles	5 EM61, 1 RTS, 1 IMU, tow-vehicle/sled, 2 vehicles	1 EM63, 1 RTS, 1 IMU, 2 vehicles
Field personnel	3 field crew & 1 escort	3 field crew & 1 escort	2 field crew & 1 escort
Processing time per acre	0.15 days	0.2 days	0.25
Processing time per anomaly	0	5 minutes	10 minutes
Detected holes excavated	100%	75%	75%

The cost comparison using the stated assumptions is provided in Table 21. For the geophysical portion of the work, a standard production survey costs around $876 per acre, compared to an EM61 discrimination survey cost of $1755 per acre and an EM63 survey cost of $4832 per acre. With a 25% reduction in false alarms, the validation costs are reduced from $11,890 per acre to $8910 per acre. This reduction is enough to make the EM61 discrimination survey more cost effective than the production survey. Due to the slow rate of survey and the length of time required for interpretation, the EM63 discrimination method is more costly. For the costs to be less than those for EM61 production, the excavations would have to be reduced by about 35% (35 anomalies per acre). In comparison, EM61 discrimination has a break-even point of 8%, or just 8 anomalies per acre. Thus according to this cost model, the discrimination efficiency of the EM61 does not have to be very high for cost savings to be realized.

Table 21. Comparison of the costs for the different modes of survey using the assumptions in Table 17 and in the bullet points immediately before that table.

System	EM61 Production			EM61 Discrimination			EM63 Discrimination		
Geophysics	**Daily Rate**	**# Days**	**Cost**	**Daily Rate**	**# Days**	**Cost**	**Daily Rate**	**# Days**	**Cost**
Mobilization/demobilization	NA	NA	$32,543	NA	NA	$34,125	NA	NA	$19,793
Survey (equipment/consumables)	$1250	10	$12,500.88	$1250	12.5	$15,626	$743	100	$74,300
Survey (labor)	$1834	10	$18,336.00	$1834	12.5	$22,920	$1375	100	$137,520
Hotel and per diem	$600	14	$8400.00	$600	17.5	$10,500	$450	140	$63,000
Processing (initial)	$463	15	$6942.00	$463	18.75	$8678	$463	25	$11,570
Processing (discrimination)	$628	0	$0.00	$628	104.167	$65,427	$628	208.33	$130,854
Deliverables/maps	$463	2	$925.60	$463	5	$2314	$463	5	$2314
Management (10%)	NA	NA	$7964.72	NA	NA	$15,959	NA	NA	$43,935
Total (Geophysics)			**$87,612**			**$175,549**			**$483,287**
Validation	**Unit Rate**	**# Holes**	**Cost**	**Unit Rate**	**# Holes**	**Cost**	**Unit Rate**	**# Holes**	**Cost**
Reacquisition	$19	10,000	$188,217.60	$19	7500	$141,163	$19	7500	$141,163
Validation	$100	10,000	$1,000,000.00	$100	7500	$750,000	$100	7500	$750,000
Total (Validation)			**$1,188,217.60**		**15,000**	**$891,163**		**15,000**	**$891,163**
TOTAL			**$1,275,829.56**			**$1,066,712**			**$1,374,450**

Obviously, the cost analysis is dependent on the values of a number of parameters that could vary widely from the numbers quoted here. These include the cost per anomaly to excavate, the number of anomalies per acre, the time required to interpret each anomaly, and the reduction in the number of excavations due to discrimination (percentage of holes to dig).

Table 22 provides a comparison of how the cost varies for different numbers of anomalies per acre and different percentages of holes to dig. As the number of anomalies per acre decreases, the discrimination efficiency of the EM63 must improve substantially for it to be cost-competitive to the production survey (due to the greater percentage cost of the geophysics compared to the validation). On the other hand, the EM61 remains cost competitive with the standard production survey with very small reductions in the number of holes to dig.

In Table 23, we investigate the influence of the time required to process each anomaly on the costs for the EM61 and EM63 surveys (assuming 100 acres and 100 anomalies/acre). As the time for processing increases, there is only a modest increase in the price. According to our price assumptions, each extra minute of interpretation time only increases the cost by $14,395. This is the equivalent of excavating an extra 144 anomalies.

Table 22. Comparison of the cost of survey for the different methods with percentages of holes to dig and different numbers of anomalies per acre.

The smallest reduction in holes to dig that produces a cost less than the standard production method is marked in green.

Number Holes	200			50		
Holes to Dig	**EM61 Prod**	**EM61 Disc**	**EM63 Disc**	**EM61 Prod**	**EM61 Disc**	**EM63 Disc**
100%	$2464K	$2624K	$3033K	$681K	$734K	$1005K
90%		$2386K	$2766K		$674K	$946K
75%		$2030K	$2410K		$585K	$857K
66%		$1816K	$2196K		$532K	$803K
50%		$1436K	$1815K		$560K	$708K
25%		$842K	$1221K		$288K	$560K

Table 23. Comparison of the cost of survey for the different methods with different percentages of holes to dig and different amounts of time required for interpretation of each anomaly.

The smallest reduction in holes to dig that produces a cost less than the standard production.

Time/Anomaly	1		5		10	
Holes to dig	**EM61**	**EM63**	**EM61**	**EM63**	**EM61**	**EM63**
100%	$1306K	$1542K	$1363K	$1600K	$1435K	$1671K
90%	$1187K	$1423K	$1244K	$1480K	$1317K	$1552K
75%	$1009K	$1244K	$1067K	$1302K	$1138K	$1374K
66%	$902K	$1138K	$958K	$1196K	$1031K	$1267K
50%	$712K	$947K	$770K	$1005K	$842K	$1077K
25%	$415K	$651K	$473K	$708K	$545K	$780K

In summary, EM61 discrimination will be more cost effective than standard EM61 production surveying if discrimination can eliminate the need to excavate four out of every 100 anomalies. If excavations can be reduced by as much as 25%, then there will be an approximate cost saving of 20%. EM61 discrimination costs compare favorably to production surveying at both higher and lower anomaly density. For a 100-acre survey with 100 anomalies/acre, an EM63 discrimination survey is likely to be more cost-effective only if it can achieve a 35% reduction in false alarms. As the anomaly density per acre decreases, this requirement increases sharply due to the need to offset the high costs of surveying with the EM63.

9.1.2 Camp Sibert Data Collection Summary

Cost categories for the Camp Sibert data collection demonstration are mobilization, field survey, data analysis, demobilization, and reporting. These costs were tracked throughout the demonstration and are presented in Table 24 (fully burdened). The data collection costs include three extra days on site to collect full-coverage data over the GPO and part of the SW area, as well as a number of transects collected to characterize the soil background of the site. The data processing costs include about 2-days implementing the RTS set-up corrections and processing the full-coverage data.

Table 24. Fully burdened costs for the Camp Sibert demonstration and the premobilization tests conducted in Ashland.

Categories	Pre-Mob Testing	Mobilization	Data Collection	Data Processing	Reporting	Total (Excluding Testing)	Total
Labor	$12,229.24	$8710.48	$21,825.19	$13,611.96	$6574.50	$50,722.14	$62,951.38
Equipment	$4238.49	$2658.44	$6054.99	$ -	$ -	$8713.42	$12,951.91
Travel	$3085.04	$3923.23	$6148.81	$ -	$ -	$10,072.04	$13,157.08
Materials and supplies	$537.94	$948.28	$16.26	$ -	$ -	$964.54	$1502.49
Std ODCs	$1068.38	$805.66	$1839.01	$561.71	$293.99	$3500.38	$4568.76
Total	**$21,159.09**	**$17,046.09**	**$35,884.27**	**$14,173.68**	**$6,868.49**	**$73,972.54**	**$95,131.62**

9.1.3 Fort McClellan Data Collection Summary

Table 25 presents the costs for preparation, data collection, processing, and analysis for the Fort McClellan demonstration. The costs to deploy the EM63 cued-interrogation system to Fort McClellan and perform the data processing included the following:

1. Predeployment testing was conducted in Ashland, OR.
2. EM63 shipping costs were incurred (the sensor was government furnished equipment).
3. No field preparation costs were incurred as this was addressed by the site manager.
4. Mobilization and demobilization costs were incurred.

5. Two field technicians were supplied by NAEVA, so although no costs were directly incurred by Sky, for purposes of reporting the complete cost of data collection, we have estimated labor and per diem travel costs for geotechnicians.

6. We surveyed 401 anomalies over a period of 16 days. The total duration of the deployment was 30 days, due to equipment problems.

7. The costs for processing and analyzing the data reflect a greater level of effort than would be projected for subsequent deployments because of the development time required for the feature extraction and classification. We made some adjustments to the underlying inversion methods and trialed a number of different inversion strategies (noise levels, masks, etc.) before settling on our final approach.

8. All costs are fully burdened.

Table 25. Fort McClellan demonstration cost summary.

Cost Element	Description	Costs
Demonstration plan	• Draft and final versions • 154 hours	$15,613
Predeployment testing	• 182 hours • Field testing equipment, prepare template, preliminary classification/discrimination algorithm development	$20,476
Instrument costs	• Equipment costs • Consumables and repairs	(Government provided EM63) $2885
Mobilization and demobilization	• Equipment shipping costs (equipment, multiple sensors shipped due to malfunctions) • Equipment packing and shipping, mobilize to site (labor, travel [Sky staff only]) • Derived from demonstration costs	$4925 $9035
Site preparation	• No unique requirements encountered	No costs incurred
Survey costs	• 300 hours Sky geophysicist, 16 hours Sky staff remote support (includes hotel and per diem costs) • Estimated NAEVA field personnel support (400 hours labor, 2 people, 20 days per diem)	$41,380 $38,000
Data processing costs	• Processing geophysicists—260 hours • Senior geophysicists, development and QC—34 hours	$31,036
Discrimination and classification	• Senior geophysicist—38 hours	$4375
Demonstration report	• Geophysicists, PI, technical edit, cost analysis—222 hours	$19,200
Total Demonstration Costs	• Actual and estimated costs for field support	**$186,925**

Table 26 presents the cost per anomaly to collect, process, and analyze the data for the 400 anomalies surveyed.

Table 26. Per anomaly cost breakdown.

Cost Element	Cost Per Anomaly
1. Data collection (based on Sky actual costs plus estimated Nova support costs)	$198
2. Data processing	$78
3. Discrimination and classification	$11
Collection, processing, discrimination and classification per anomaly costs (sum of items 1 through 3)	$287
Complete per anomaly costs (including predeployment testing, work plan and report preparation—total demonstration costs/400 anomalies)	$467

9.1.4 Projected Costs for Future EM63 Cued-Interrogation Deployments

Because both the Camp Sibert and Fort McClellan surveys were conducted as demonstrations with greater development and analysis effort than would be expected to deploy the system in a live site survey, we developed assumptions and costs that might be logically associated with future deployments.

For estimating the costs (fully burdened) to deploy this system to other sites (Table 27), we made the following assumptions:

- Costs for preparing a site-specific work plan are assumed to be less than that for either the Camp Sibert or Fort McClellan efforts because the effort would be planned and executed as a production survey rather than a technology demonstration.
- Mobilization/demobilization
  - The survey requires a 1200 mile mobilization from Denver (assumes a site in the southeastern United States).
  - Mobilization includes 1 day preparation, 1 day travel, 1 day setup on site and 1 day for a test plot survey (and associated processing) for 1 field geophysicist and 2 geotechnicians.
  - Demobilization includes 1 day of packing up on site and 1 day of organization back at home base.
- 400 anomalies will be measured, 25 anomalies per day (16 days to complete data collection, plus weekends).
- Daily rates are fully burdened and assume 10 hour workdays.
- Processing requires a Geophysicist II level analyst while interpretation and QC is performed at the Geophysicist V level.
- Per diem and hotel are assumed to total $150 per person per day, with 1.4 days of per diem/hotel per day in the field (5-day week, 2 days on the weekend) and vehicles (assumed 1 rental vehicle required).

- Equipment charge includes all equipment, consumables, and supplies.
- There is a cost of approximately 15% for report preparation, administration, and reporting support.

Table 27. Future deployment cost estimates.

Cost Element	Description	Costs
Predeployment planning and work plan	• Site coordination and logistics • Site-specific work plan	$9130
Mobilization and demobilization	• Equipment shipping costs ($1000) • Field team travel to and from the location, 1 day setup, 1 day test plot survey, 1 day packing/shipping upon survey completion	$16,897
Site preparation	• Assume site manager will address	No costs incurred
Survey costs	• Field geophysicist and 2 geotechnicians • 400 anomalies, 25 anomalies per day • 16 days data collection • Field team per diem, including weekends	$60,118
Data processing and analysis costs	• Processing—Geophysicist II, QC; Geophysicist V (assumed level of effort 5% of processing) • **20 minutes per anomaly**	$11,294
	• Processing—Geophysicist II, QC; Geophysicist V (assumed level of effort 5% of processing) • **10 minutes per anomaly**	$5639
	• Processing—Geophysicist II, QC; Geophysicist V (assumed level of effort 5% of processing) • **5 minutes per anomaly**	$2897
Demonstration report	• Reporting and support	$19,200
Total Estimated Demonstration Costs	• Total costs, assuming 20 minutes per anomaly • 10 minutes/anomaly • 5 minutes/anomaly	**$116,639** **$110,984** **$108,242**

The projected costs for future deployments are substantially less than the actual costs to conduct the Fort McClellan data collection and analysis. The lower costs reflect assumptions that there would be fewer days in the field required to address equipment issues (this extended the Fort McClellan field demonstration by more than a week), and the fact that there was a more significant effort required to develop the appropriate data processing and discrimination approaches. We assume that the lessons learned from this effort will benefit future deployments and reduce the overall costs to perform similar surveys.

9.2 CAMP SIBERT DISCRIMINATION

The demonstration costs for applying discrimination techniques to the data collected using each of the different sensor systems were tracked throughout the demonstration (Table 28). Only approximate costs and times were available for the one UBC employee (LinPing Song) who participated in the data analysis component of the project. The magnetometer interpretation

required less time and cost than the other full-coverage methods. The MTADS EM61 required less time to interpret than the contractor EM61, partly because the EM61 data required more time to organize and manage (due to the way it was delivered to demonstration participants). For cooperative inversion, we list two costs. The first represents the additional costs associated with inversion, QC, and interpretation using the magnetometer locations as constraints. However, these are not a good indication of the true cost of cooperative inversion as the anomalies had already been preprocessed and masked prior to inversion. To estimate the cost of cooperative inversion we add the cost of the magnetometer interpretation (calculated as the number of EM anomalies times the cost per anomaly for magnetometer interpretation) to the cost of the single inversion interpretation of the EM data. These numbers are presented in parentheses.

For the cooperative inversion dig sheets, the first number represents the additional cost of cooperative inversion (after single inversions have been complete). The second number in number in parentheses represents the estimated total cost of cooperative inversion (including the magnetometer interpretation).

Table 28. Camp Sibert discrimination cost summary.

Category	Hours	Cost	# Anomalies	Cost per Anomaly
Pre-demo testing prep	542	$62,652		
MTADS EM61 interpretation	165	$15,259	870	$17.54
Magnetometer interpretation	82	$8167	969	$8.43
EM63 interpretation	65	$6431	178	$36.13
Contractor EM61 interpretation	186	$18,685	633	$29.52
MTADS EM61 cooperative	46	$4966 ($22,593)	870	$5.71 ($25.97)
EM63 cooperative	16	$1491 ($6755)	178	$8.38 ($37.95)
Demonstration report	228	$38,646		
Total	**1330**	**$156,297**		

10.0 IMPLEMENTATION ISSUES

A comprehensive discussion of general implementation issues that need to be considered before implementation of discrimination methods at DoD sites can be found in Nelson et al. (2008). Specific implementation issues related to the technologies demonstrated here are:

- All major hardware components (sensors, positioning systems) are COTS, with some of the carts custom made and fabricated at Sky.
- UXOLab software was used for all the processing and interpretation activities described in the report. While UXOLab is available under license from UBC, it is not suitable for general distribution to government contractors. Firstly, using the software successfully requires advanced knowledge of geophysical inversion and statistical classification. Secondly, while the software doesn't require the user to have a Matlab license, it was built entirely within the Matlab software environment to support the needs of UXO researchers. Thirdly, UBC is not set up to provide maintenance and support for the software. Sky is currently in the process of porting the Matlab code to C++ so that the software will be accessible with Geosoft Oasis Montaj (the industry standard processing software).
- Access to the appropriate software and hardware is the first critical step in successful implementation of the technology. Ensuring that operators have the appropriate level of training and experience is the second. This requirement for training and experience applies to both the feature vector extraction and classification steps. While the process of feature extraction is straightforward and reliable for a large number of anomalies, there is always a certain percentage of anomalies where the default parameters don't produce a sensible or reliable result. Recognizing those situations is critical to minimize the probability of a false negative, while being able to appropriately adjust the strategy is important for reducing the FP count. For classification, each of the sites essentially used a common strategy of a combination of size- and time-based feature vectors. While this strategy was successful for the mix of ordnance and clutter encountered at these sites, it may not always be the most effective and reliable strategy. Thus, the analyst must be capable of exploring (and understanding the potential pitfalls of) different feature spaces that could be used for discrimination. In summary, we believe that a very solid understanding of both feature extraction and classification are required for the discrimination technology to be reliably applied to different DoD sites.

This page left blank intentionally.

11.0 REFERENCES

Billings, S.D., L.R. Pasion, and D.W. Oldenburg. 2002. Discrimination and identification of UXO by geophysical inversion. Phase II: Inversion of total-field magnetics, *USACE Engineer Research and Development Center,* Tech. Rep. ERDC/GSL TR-02-16.

Billings, S.D. 2004. Discrimination and classification of buried unexploded ordnance using magnetometry: *IEEE Transactions of Geoscience and Remote Sensing*, 42, 1241-1251.

Billings, S.D., L.R. Pasion, J. Jacobson, S. Kinsgbury, D. Oldenburg, L. Beran, L.P. Song, D. Sinex, and N. Lhomme. 2007. Demonstration Report for the Former Lowry Bombing and Gunnery Range, Project 200504: Practical Discrimination Strategies for Application to Live Sites, Demonstration Report. ESTCP Program Office.

Billings, S.D. 2008. Data Modeling, Feature Extraction, and Classification of Magnetic and EMI Data, ESTCP Discrimination Study, Camp Sibert, AL, Project 200504: Practical Discrimination Strategies for Application to Live Sites, Demonstration Report. ESTCP Program Office.

Billings, S.D. and K. Kingdon. 2009. Discrimination Using the Geonics EM63 in a Cued Interrogation Mode at Fort McClellan, AL, Project 200504: Practical Discrimination Strategies for Application to Live Sites, Demonstration Report, ESTCP Program Office.

Collins, L., Y. Zhang, J. Li, H. Wang, L. Carin, S. Hart, S. Rose-Phersson, H. Nelson, and J.R. McDonald. 2001. A comparison of the performance of statistical and fuzzy algorithms for unexploded ordnance detection: *IEEE Transactions on Fuzzy Systems*, 9, 17-30.

Carin, L., Y. Zhang, and X. Liao. 2004. Detection of Buried UXO via Active Selection of Labeled Data. Proceedings from the UXO Forum, St. Louis, March 9-12, 2004. ESTCP, 2006, ESTCP Discrimination Study Demonstration Plan, Draft 2, October.

Hart, S. J., R. E. Shaffer, S. L. Rose-Pehrsson, and J. R. McDonald. 2001. Using Physics Based Modeler Outputs to Train Probabilistic Neural Networks for Unexploded Ordnance (UXO) Classification in Magnetometry Surveys. *IEEE Trans. Geosci. Remote Sensing* 39, 797-804.

Lhomme, N., D. Oldenburg, L. Pasion, D. Sinex, and S. Billings. 2008. Assessing the Quality of Electromagnetic Data for the Discrimination of UXO Using Figures of Merit, *Journ. of Envir. Engin. and Geophys. (JEEG)*, September.

Nelson, H.H., T.H. Bell, J.R. McDonald, and B. Barrow. 2003. Advanced MTADS Classification for Detection and Discrimination of UXO, Final Report, MM-4003, January.

Nelson, H.H., K. Kaye, and A. Andrews. 2008. ESTCP Pilot Program classification approaches in munitions response: ESTCP Program Office.

Pasion, L., and D. Oldenburg. 2001. A Discrimination Algorithm for UXO Using Time Domain electromagnetics*: Journal of Engineering and Environmental Geophysics*, 28, 91-102.

Pasion, L.R., S.D. Billings, and D.W. Oldenburg. 2003. Joint and Cooperative Inversion of Magnetics and Electromagnetic Data for the Characterization of UXO Discrimination Problems: *The Symposium on the Application Geophysics to Engineering and Environmental Problems (SAGEEP), San Antonio*.

Zhang, Y., L.M. Collins, and L. Carin. 2003a. Model-based statistical signal processing for UXO discrimination: Performance results from the JPG-V demonstration: *Proceedings of SPIE Volume 5089*, 1116-1126.

Zhang, Y., L. Collins, H. Yu, C.E. Baum, and L. Carin. 2003b. Sensing of Unexploded Ordnance with Magnetometer and Induction Data: Theory and Signal Processing. *IEEE Trans. Geosci. Remote Sensing,* 41, 1005-1015.

APPENDIX A

POINTS OF CONTACT

Point of Contact	Organization	Phone Fax E-Mail	Role
Stephen Billings	Sky Research, Inc. 112A East Mall Vancouver, BC V6T 1Z3	541-552-5185 604-827-3221 stephen.billings@skyresearch.com	Principal Investigator
Dr. Len Pasion	Sky Research, Inc. 112A East Mall Vancouver, BC V6T 1Z3	541-552-5186 604-827-3221 len.pasion@skyresearch.com	Data Analyst, QC
Kevin Kingdon	Sky Research, Inc. 112A East Mall Vancouver, BC V6T 1Z3	541-552-5188 604-827-3221 kevin.kingdon@skyresearch.com	Field Geophysicist, Data Analyst
Dr. Lin Ping Song	Department of Earth and Ocean Sciences University of British Columbia Vancouver, BC V6T 1Z4	604-822-1819 604-822-6088 lpsong@eos.ubc.ca	Data Analyst
Dr. Douglas Oldenburg	Department of Earth and Ocean Sciences University of British Columbia Vancouver, BC V6T 1Z4	604-822-5406 604-822-6088 doug@eos.ubc.ca	Key Technical Support, Director, UBC-GIF
Dr. Anne Andrews	SERDP/ESTCP 901 N. Stuart Street, Suite 303 Arlington, VA 22203	703-696-3826 703-696-2114 anne.andrews@osd.mil	Program Manager, Munitions Management
Dr. Herb Nelson	SERDP/ESTCP 901 N. Stuart Street, Suite 303 Arlington, VA 22203	703-696-8726 703.696.2114 Herbert.nelson@osd.mil	Program Manager, Munitions Management

This page left blank intentionally.

ESTCP Program Office

901 North Stuart Street
Suite 303
Arlington, Virginia 22203

(703) 696-2117 (Phone)
(703) 696-2114 (Fax)

E-mail: estcp@estcp.org
www.estcp.org