FINAL REPORT

Analysis of EMI and Magnetic data; Aberdeen Proving Ground

ESTCP Project MM-0210:
Feature-based UXO Detection and Discrimination

December 2008

Dean Keiswetter
Science Applications International Corporation
120 Quade Dr.
Cary, NC 27513

Approved for public release; distribution unlimited.

Environmental Security Technology Certification Program

Table of Contents

Table of Contents i
Lists of Figures iii
Lists of Tables v
Acknowledgements v
List of Acronyms vi
Abstract 1
1 Introduction 1
1.1 Background 1
1.2 Objectives of the Demonstration 2
1.3 Regulatory Drivers 2
1.4 Stakeholder/End-User Issues 3
2 Technology Description 3
2.1 Technology Developments and Application 3
2.1.1 UX-Analyze 3
2.1.2 Magnetic Array Data 4
2.1.3 EM61 Cart Data 4
2.1.4 EMI Template Data 5
2.1.5 EM61 Array Data 6
2.2 Previous Testing of the Discrimination Technology 7
2.2.1 UX-Analyze 7
2.2.2 Characterization Modules 7
2.2.3 Classification Modules 11
2.2.4 Data Analysis Documentation 13
2.3 Factors Affecting Cost and Performance 14
2.4 Advantages and Limitations of the Technology 15
3 Demonstration Design 16
3.1 Performance Objectives 16
3.2 Selecting Test Sites 16
3.3 Test Site History/Characteristics 16
3.4 Present Operations 18
3.5 Pre-Demonstration Testing and Analysis 18
3.6 Testing and Evaluation Plan 19
3.6.1 Demonstration Set-up and Start-up 19
3.6.2 Period of Operation 19
3.6.3 Area characterized or Remediated 19
4 Performance Assessment 22
4.1 Performance Criteria 22
4.2 Performance Confirmation Methods 23
4.3 Data Analysis, Characterization, and Evaluation 24
4.3.1 Cued Target Collection Comparisons and Results 24
4.3.2 Blind Test Scores 36
4.3.3 Effects of Changing the Labeled Data and Feature Selection 39

4.3.4 Analysis Time 40
4.3.5 Qualitative Metrics 40
4.4 Discussion 41
5 Cost Assessment 42
5.1 Cost Reporting 42
6 Implementation Issues 42
6.1 End-Users Issues 42
7 References 43
8 Points of Contact 44
ESTCP 44
SAIC 44

Lists of Figures

Figure 2-1. Photograph of G-tek's magnetic sensor system during data acquisition at the open field area of APG. 4
Figure 2-2. Photograph of TTFWs EM61 MK2 sensor system during data acquisition at the open field area of APG. 5
Figure 2-3 Photograph taken during collection of EM61 MK2data using a template at APG. 6
Figure 2-4. Photograph of NRLs Multi-Sensor Towed Array Detection Systems (MTADS). 7
Figure 2-5. Screen snapshots showing UX-Analyze during target selection. 8
Figure 2-6. Screen snapshots showing the user interface during data inversion. 8
Figure 2-7. Top – schematic showing the inclination and declination of the source objects. The inclination and declination range from 0 to 90 degrees in 15-degree increments. Bottom – synthesized electromagnetic and magnetic data, left and right respectively. 9
Figure 2-8. Screen snapshots showing from left to right along each row (i) the segmentation of individual anomalies, (ii) the graphical interface used to identify and select individual anomalies, and (iii) anomaly-specific comparisons of measured and modeled data. Row (a) shows images for UX-Analyze EMI data, (b) DAS EMI, (c) UX-Analyze magnetic, and (d) DAS magnetic. 10
Figure 2-9. Screen snapshots showing the GLRT classification dialogue in UX-Analyze. 13
Figure 2-10. UX-Analyze generates a one-page summary for each anomaly. In the anomaly summaries shown above, the measured data is shown in the upper left hand corner, the inverted model parameters in the middle left, the forward model in the upper rights, and a profile in the lower left corner. EMI data for the anomaly are shown in the left summary, and magnetic data on the right. 14
Figure 3-1. Aerial photograph of the Aberdeen Proving Ground Test Site. 18
Figure 3-2. Color contour map of the EM61 Mk2 data (0.366ms time gate) collected by TTFW in November 2003. The symbols represent targets for cued investigation. 20
Figure 3-3. False color images showing G-tek magnetic (left) and TTFW EMI (right) data from the Open Field area. The symbols identify anomalies included in the analysis. 21
Figure 4-1. SNR versus dipole fit errors for all targets surveyed using the template. 24
Figure 4-2. Actual versus fitted depth estimates, UXO only. 25
Figure 4-3. Depth errors for UXO 26
Figure 4-4. Polarization plots for labeled UXO. 27
Figure 4-5. Discrimination performance for EM61 Cart data (cued target collection). A) ROC curve B) Comparison of all intrinsic features versus selected C) table listing selected and intrinsic features (names ending in "_R" or "_sum" are a ratio or sum respectively) 29
Figure 4-6. Discrimination performance for EM61 Array data (cued target collection). A) ROC curve B) Comparison of all intrinsic features versus selected C) table listing selected and intrinsic features (names ending in "_R" or "_sum" are a ratio or sum respectively) 30
Figure 4-7. Discrimination performance for EM61 Template data (cued target collection). A) ROC curve B) Comparison of all intrinsic features versus selected C) table listing selected and intrinsic features (names ending in "_R" or "_sum" are a ratio or sum respectively) ... 31
Figure 4-8. Polarization plots segmented by survey approach and by size-based clusters of UXO. 32

Figure 4-9. Discrimination performance for EM61 Template data, restricting the TOI subclass to UXO smaller than 60mm. A) ROC curve B) Comparison of all intrinsic features versus selected C) table listing selected and intrinsic features (names ending in "_R" or "_sum" are a ratio or sum respectively). 33
Figure 4-10 Discrimination performance for EM61 Template data, restricting the TOI subclass to UXO larger than 57mm but smaller than 105mm projectiles. A) ROC curve B) Comparison of all intrinsic features versus selected C) table listing selected and intrinsic features (names ending in "_R" or "_sum" are a ratio or sum respectively) 34
Figure 4-11. Discrimination performance for EM61 Template data, restricting the TOI subclass to UXO larger than 105mm. A) ROC curve B) Comparison of all intrinsic features versus selected C) table listing selected and intrinsic features (names ending in "_R" or "_sum" are a ratio or sum respectively) 35
Figure 4-12. A false-color image of the EM61 Cart data, calibration lanes, is shown on the left. Fit statistics for targets in the calibration lanes are shown on right. 36
Figure 4-13. A false-color image of the magnetic data, calibration lanes, is shown on the left. Fit statistics for targets in the calibration lanes are shown on right 37
Figure A-1. Discrimination performance for Mag data (Open Field supplemental labels only). A) ROC curve B) Comparison of all intrinsic features versus selected C) table listing selected and intrinsic features. 45
Figure A-2. Discrimination performance for EMI Cart data (Open Field supplemental labels only). A) ROC curve B) Comparison of all intrinsic features versus selected C) table listing selected and intrinsic features. 46

Lists of Tables

Table 3-1. Qualitative Performance Objectives 16
Table 3-2. Inert ordnance emplaced at APG. 17
Table 4-1. Criteria for this Demonstration 22
Table 4-2. Quantitative Performance Objectives: Open Field Blind Test (MAG GLRT) 23
Table 4-3. Quantitative Performance Objectives: Open Field Blind Test (EMI GLRT) 23
Table 4-4. SNR and Fit Error statistics across data sets 24
Table 4-5. Selected features used by the classifiers for blind testing 37
Table 4-6. EMI Blind Scoring Results 38
Table 4-7. Magnetic Blind Scoring Results 38
Table 4-8. Depth and XY Location Error Statistics 38
Table 4-9. TOI Retention Rates as a function of dipole error 38
Table 4-10. Non-TOI Rejection Rate as a function of dipole error 39
Table 4-11. Depth Errors (actual minus estimated, meters) 39
Table 4-12. XY Distance Errors (meters) 39
Table 5-1. Cost categories and details 42

Acknowledgements

Contributors to this demonstration include Science Applications International Corporation (SAIC; formerly AETC Incorporated (lead)) and Duke University under funding from the Environmental Security Technology Certification Program Office.

List of Acronyms

ACRONYM	DEFINITION
APG	Aberdeen Proving Ground
BRAC	Base Realignment and Closure
COTS	Commercial-of-the-Shelf
DoD	Department of Defense
EMI	Electromagnetic induction
EPA	Environmental Protection Agency
ESTCP	Environmental Security Technology Program
FUDS	Formerly Used Defense Sites
GLRT	Generalized likelihood ration test
GPS	Global Positioning System
GX	Geosoft Executables
IDL	Interactive Data Language
KNN	k-nearest neighbor classifier
HASP	Health and Safety Plan
ROC	Receiver operating characteristic
RVM	Relevance vector machine
SERDP	Strategic Environmental Research and Development Program
SNR	Signal to Noise Ratio
SVM	Support vector machine
TTFW	Tetra Tech – Foster Wheeler
UXO	Unexploded Ordnance

Abstract

Buried Unexploded Ordnance (UXO) is a serious and prevalent environmental problem facing Department of Defense (DoD) facility managers. Not limited to active sites and test ranges, these problems also occur at DoD sites that are currently inactive and in areas adjacent to military ranges that belong to the civilian sector or are under control of other government agencies. The exact amount of land affected is uncertain, but it is generally agreed to be in excess of 10 million acres in the continental United States. Estimated cleanup costs exceed 10 billion dollars. Over the past decade, SERDP and ESTCP have invested heavily in developing survey data analysis and processing techniques for use with commercial and custom sensor suites that improve the detection, characterization, and classification of UXO and anthropic clutter. These techniques include procedures for quantitatively characterizing and classifying the anomalies.

In this report, we compare and contrast the performances of feature-based discrimination decisions on magnetic and EMI data acquired at the Aberdeen Proving Ground (APG) Standardize Test Site. The APG test site was selected because it was designed to imitate conditions commonly found at munitions response sites, and includes emplaced ordnance ranging from 20mm projectiles to general practice bombs.

We analyzed magnetic and electromagnetic induction data that was acquired using standard production methodologies by others. Two EM61 MK2 data sets, one acquired using an array and the other acquired using a commercial cart configuration, were included in our analysis. To bound performance expectations, we also acquired and analyzed EM61 MK2 EMI data collected using a template. The template data provide an upper bound on performance because the deleterious effects of motion noise and positioning errors are mitigated.

The anomaly characterization algorithms utilized in this demonstration assume a dipolar source and derive the best set of induced dipole model parameters that account for the spatial variation of the signal as the sensor is moved over the object. The model parameters are location, depth, three polarization coefficients (EMI only), magnetic moment (magnetic only), and three orientation angles. The primary metrics chosen to report results include the target of interest (TOI) retention rate and non-TOI rejection rates.

Blind test results showed that, for the production survey EMI data and all UXO, it was possible to correctly identify 82% of the UXO, while correctly rejecting 25% of the clutter items. Blind test results for the magnetic data correctly classified 87% of the UXO while correctly rejecting 27% of the emplaced clutter.

Discriminating UXO from clutter based upon their inverted polarizabilities is inherently limited by how well the measured data can be replicated using a dipole model source and by the extent to which the class polarization distributions overlap in feature space. The mean dipole fit error for the template EMI data was 4.7% (standard deviation of 5.1), which is quite good, and allows us to use the template data to evaluate the polarization distributions. By comparison, the mean fit error

the production cart-mounted EMI data was 24.1% (standard deviation of 17.2). Using the fitted parameters from the template EMI data to evaluate the separability of the TOI and non-TOI clusters, we find that the polarizabilities from non-TOI objects overlap significantly in feature space with those derived from the larger UXO. Thus, even with well positioned static data, discrimination performance at this site is limited by the distributions of the TOIs relative to the non-TOI objects. Interestingly, however, we also find that subclasses of UXO, namely 57mm and smaller UXO, can be effectively discriminated at this site. This finding is probably unique to APG and should not be extrapolated to live sites that possess large numbers of small fragments and clutter items.

1 Introduction

1.1 Background

Buried Unexploded Ordnance (UXO) is arguably one of the most serious and prevalent environmental problems currently facing Department of Defense (DoD) facility managers. Not limited to active sites and test ranges, these problems also occur at DoD sites that are currently inactive and in areas adjacent to military ranges that belong to the civilian sector or are under control of other government agencies. The exact amount of land affected is uncertain, but it is generally agreed to be in excess of 10 million acres in the continental United States. UXO mitigation and remediation requirements assume even more compelling proportions when the DoD lands involve Formerly Used Defense Sites (FUDS) or Base Realignment and Closure (BRAC) sites. These sites must be certified as suitable for the end use depending on the pending disposition. Oversight and evaluation of these processes involve non-DoD agencies including the Environmental Protection Agency (EPA); state, county, and local governments; and the civilian community.

An objective of this project is to make available and document the capabilities of feature-based discrimination techniques. Specific goals include transitioning physics-based characterization and classification algorithms for magnetic and electromagnetic data to a commercial product, and conducting demonstrations at live sites with the aim of discriminating targets of interest from targets that are not of interest in order to mitigate risk during the recovery process.

Feature-based characterization and classification schemes have improved discrimination performance in some demonstrations (Robitaille et al., 1999). These algorithms, however, are not readily available to the user community and have had limited exposure to data acquired under 'production-imposed' constraints. ESTCP project MM-0210 is designed to address the availability problem as well as further test the ability to improve decisions using features - or model parameters - derived from field data.

Our technical approach promotes the selection of potential UXO targets using quantitative evaluation criteria and transparent decision-making processes. As such, we developed an analysis framework within Oasis montaj and integrated previously developed, physics-based characterization and statistical classification algorithms. The analysis algorithms provide quantitative evaluation criteria. Transparency is achieved by leveraging the professional, flexible, and visual computing environment inherent in Oasis montaj. Oasis montaj is a geophysical data processing and visualization package developed and marketed by Geosoft Incorporated. It has a large capacity database, a professional graphic interface, and an established client base. Oasis montaj was selected in order to leverage its significant capabilities, marketing channels, and customer support services.

1.2 Objectives of the Demonstration

The objective of this demonstration was to shakedown UX-Analyze and to document, contrast, and compare the performance of feature-based discrimination decisions on multiple magnetic and EMI data acquired at the Aberdeen Proving Ground Standardize Test Site.

1.3 Regulatory Drivers

The Senate Report (Report 106-50), pages 291–293, accompanying the *National Defense Authorization Act for Fiscal Year 2000* (Public Law 106-65), included a provision entitled "Research and development to support unexploded ordnance clearance, active range unexploded ordnance clearance, and explosive ordnance disposal." This provision requires the Secretary of Defense to submit to the Congressional defense committees a report that gives a complete estimate of the current and projected costs, to include funding shortfalls, for UXO response at active facilities, installations subject to BRAC sites and FUDS.

The following statements are taken verbatim out of the DoDs 2001 Report to Congress:

> *"Decades of military training, exercises, and testing of weapons systems has required that we begin to focus our response on the challenges of UXO. Land acreage potentially containing UXO has grown to include active military sites and land transferring or transferred for private use, such as BRAC sites and FUDS. DoD responsibilities include protecting personnel and the public from explosive safety hazards; UXO site cleanup project management; ensuring compliance with federal, state, and local laws and environmental regulations; assumption of liability; and appropriate interactions with the public.*
>
> *…Through limited experience gained in executing these activities, it has become increasingly clear that the full size and extent of the impact of sites containing UXO is yet to be realized. … DoD has completed an initial baseline estimate for UXO remediation cost. This report provides a UXO response estimate in a range between $106.9 billion and $391 billion in current year [2001] dollars. …Technology discovery, development, and commercialization offers some hope that the cost range can be decreased. …*
>
> *… **Objective**: **Develop standards and protocols** for navigation, geo-location, data acquisition and **processing**, and performance of UXO technologies.*
> - *Standard, high quality archived data are needed for optimal data processing of geophysical data, re-acquisition for response activities, quality assurance, quality control, and review by all stakeholders. In addition standards and protocols are required for evaluating UXO technology performance to aid in selecting the most effective technologies for individual sites.*
> - ***Standard software and visualization tools are needed to provide regulatory and public visibility to and understanding of the analysis and decision process made in response activities**."*

1.4 Stakeholder/End-User Issues

The stakeholders and end-users of this data processing and analysis technology include private contractors that conduct geophysical investigations in support of UXO clean up programs and governmental employees that provide technical oversight. This demonstration introduced the stakeholders and end-users to data products associated with this analysis approach and to the inherent transparency of the decision-making process.

2 Technology Description

2.1 Technology Developments and Application

All field production data were processed using Oasis montaj and analyzed using UX-Analyze. The field production datasets include magnetic data acquired by G-tek and electromagnetic induction (EMI) data acquired by Tetra Tech-Foster Wheeler (TTFW). These data were acquired and scored as part of ESTCP project MM-0103 entitled "Standardized UXO Technology Demonstration Sites" (http://aec.army.mil/usaec/technology/uxo03.html). Summaries of the sensing systems are provided below.

To bound performance expectations, we also acquired and analyzed EM61 MK2 EMI data collected using a template. The template data provide an upper bound on performance because the deleterious effects of motion noise and positioning errors are mitigated.

2.1.1 UX-Analyze

The anomaly characterization algorithms developed during the past decade assume a dipolar source and derive the best set of induced dipole model parameters that account for the spatial variation of the signal as the sensor is moved over the object. The model parameters are target location and depth, three dipole response coefficients corresponding to the principle axes of the target (EMI only), magnetic moment (magnetic only), and the three angles that describe the orientation of the target. The size of the target can be estimated using empirical relationships between either the dipole moment for magnetic data or the sum of the targets' response coefficients. Cylindrical objects, like most UXO, have one large coefficient and two smaller, equal coefficients. Plate-like objects nominally have two large and one small coefficient.

UX-Analyze was developed during the first year of the program to facilitate efficient UXO data analysis within the Oasis montaj environment. It consists of multiple databases, custom graphical interfaces, and data visualizations. UX-Analyze provides the infrastructure to systematically identify and extract anomalies from the dataset, call the characterization routines, store the fitted source parameters for each anomaly, perform target classification, and document the analysis. Once the analysis is complete, individual images for each anomaly can be automatically produced for documentation purposes.

2.1.2 Magnetic Array

G-tek deployed a hand-held TM-4 magnetometer array and a digital global positioning system (DGPS) (Figure 2-1). The TM-4 is a self-contained magnetometer system, which was configured with four, optically pumped magnetic sensors mounted in an array oriented perpendicular to the survey direction. The sensor separation is 30 cm and ground clearance 25 cm. The TM-4 was interfaced with a real-time kinematic DGPS. The sensor data was positioned in post processing to achieve a reported position accuracy of five cm or less.

A two-person crew operated the TM-4 system. One-person carries the sensor array to which is attached the DGPS antenna and odometer system. A five-meter cable eliminating interference at the sensors from the other hardware separates the two operators. Data processing consists of magnetic base-station subtraction, optional band-pass spatial filtering to enhance particular source depths, grading and imaging. G-tek surveyed the open field area in 57 man hours.

Figure 2-1. Photograph of G-tek's magnetic sensor system during data acquisition at the open field area of APG.

2.1.3 EM61 Cart

TTFW deployed a Geonics EM61 MK2 time-domain geophysical sensor and a Leica Series 1100 Robotic Total Station (RTS) laser positioning systems at APG (Figure 2-2). The EM61 MK2 consists of two coils 100 by 100 cm that are oriented in a horizontal coplanar fashion and separated by a vertical distance of 40 cm. The lower coil was maintained 40cm above the ground by the nonmagnetic wheel assemblage provided by the manufacturer. The Leica Series 1100 RTS consists of a laser-based total station survey instrument (transmitter), prism (receiver), and RCS 100 remote control. The receiver prism is centered over the EM61 MK2 coils, and the RTS automatically tracks the prism at distances of several thousand feet to an accuracy of

approximately one inch. Transects were spaced no more than 2 to 2.5 feet apart in order to detect the smaller objects.

The positioning and EM61 MK2 signal data were merged with the software developed by TTFW. The data were leveled (background subtraction as determined by mode of data) during processing and are output as an ASCII file that contained the state planar coordinates of each measurement location in feet, EM61 MK2 signal intensity for each time gate in millivolts, and a quality identifier for each recorded position (number 1-6, based on standard deviation). TTFW surveyed the open field area in 64 man hours.

Figure 2-2. Photograph of TTFWs EM61 MK2 sensor system during data acquisition at the open field area of APG.

2.1.4 EMI Template

The template data collection used an EM61 MK2 with a 1.0x0.5 meter coil. We recorded three bottom coil time gates (0.216ms, 0.366ms and 0.660ms) and one upper coil time gate (0.660ms). The data were collected on a 7x5 point rectangle grid (nodes separated by 25cm and 50cm respectively resulting in a 1.5x2.0m area). The grid was elevated 20 cm above the ground, and the EM61 was positioned directly on the grid, without wheels (Figure 2-3). Sensor location on the grid was precisely controlled by lining up cross hairs on the EM61 bottom coil with grid line intersections. The choice of grid layout and elevation was based on results of simulations of data inversion sensitivity as well as past experience.

Previous experience with test stand data and data collected at various sites indicated that the grid does not need to be precisely centered over the target in order to recover useful inversion results. Offsets of approximately 40 cm between the actual center of the target and the center of the grid template produce good inversion results. We performed the collections within these guidelines.

The data collection per target required approximately 20 minutes. This reflects the time required to align the sensor and collect a couple of seconds of data at each grid point, as well as recording background levels before and after the grid.

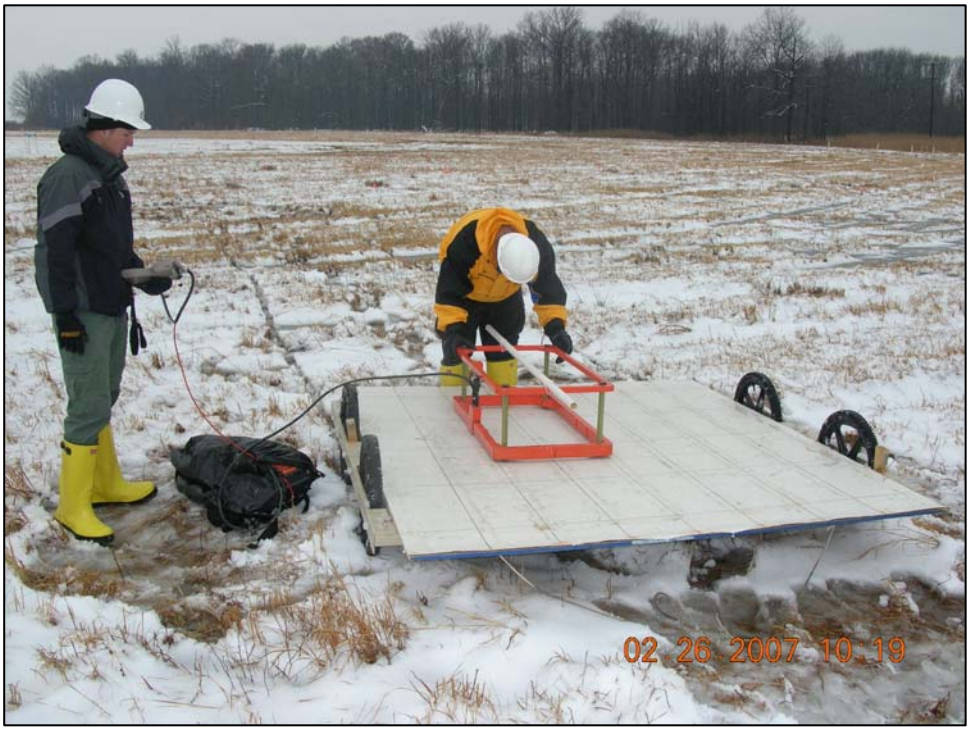


Figure 2-3 Photograph taken during collection of EM61 MK2data using a template at APG.

2.1.5 EM61 Array

The Multi-Sensor Towed Array Detection Systems (MTADS) hardware consists of a low-magnetic signature vehicle that is used to a 3-sensor array of EM61 MK2 sensors or eight cesium vapor magnetometers (Figure 2-4). The pulsed-induction sensors (specially modified Geonics EM61 MK2's) are deployed as an overlapping array of three sensors.

The sensor positions are measured in real-time (five Hz) using the latest real time kinematic global positioning system technology. The EM61 data are recorded in a down-track sampling of approximately 15 cm and a cross-track interval of 50 cm. In order to obtain sufficient views of the targets, NRL collects data in two orthogonal surveys. NRL surveyed the accessible portions of the open field area in 7 hours.

Figure 2-4. Photograph of NRLs Multi-Sensor Towed Array Detection Systems (MTADS).

2.2 Previous Testing of the Discrimination Technology

We performed preliminary tests of the demonstration data analysis and classification technology during algorithm development. A number of the user graphical interfaces and dialogue boxes are briefly described below. Each of these capabilities was tested on small data sets prior to the demonstration.

2.2.1 UX-Analyze

UX-Analyze allows users to systematically identify, extract, edit, and store data around individual anomalies. It provides efficient data structures and access for the analysis algorithms, stores the fitted parameters, and allows for multiple data types and surveys. This module is the interface between Oasis montaj and the demonstration analysis software (Figure 2-5).

2.2.2 Characterization Modules

Characterization routines for magnetic and EMI data have been integrated with UX-Analyze framework. These 3-D routines include graphic displays and controls that allow the user to manually select and filter the input data for each anomaly (Figure 2-6). The derived model parameters are stored in a master target database. The characterization modules, or inversion routines, were previously developed by AETC Incorporated (SAIC acquired AETC in 2006) for the *MTADS* Data Analysis System under funding from ESTCP and SERDP. The MTADS DAS codes were prototyped using the Interactive Development Language.

Algorithm equivalency tests verified that the C-based inversion routines embedded in Oasis montaj produce identical performances as the original formulations. Magnetic and EMI data were synthesized for forty-nine sources that have a unique combination of inclination and declination but constant moment and depth of burial. The layout and noise-free synthetic data are shown in Figure 2-7 while Figure 2-8 shows screen snapshots of the user interfaces at various

stages of the analysis process. To ensure that each routine received the exact same input for each anomaly, we extracted data samples around each anomaly individually once and then used the extracted data subsets as input for both. Source parameters were then calculated using the two inversion routines and compared. No statistically significant differences were observed.

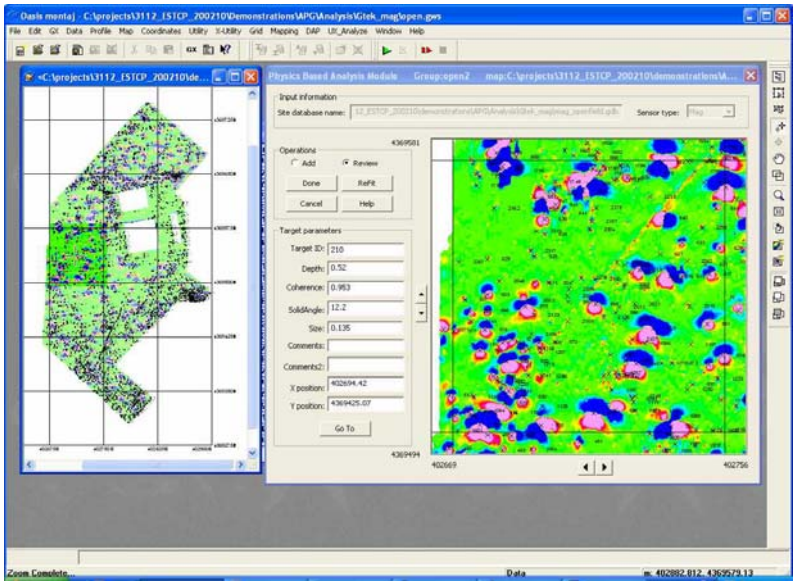


Figure 2-5. Screen snapshots showing UX-Analyze during target selection.

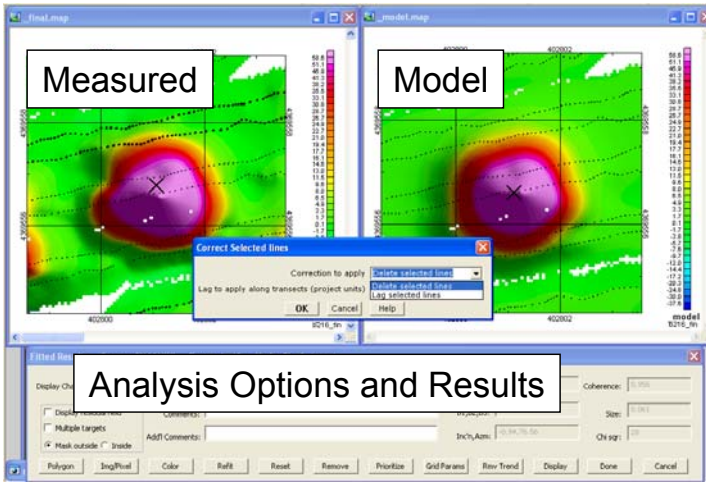


Figure 2-6. Screen snapshots showing the user interface during data inversion.

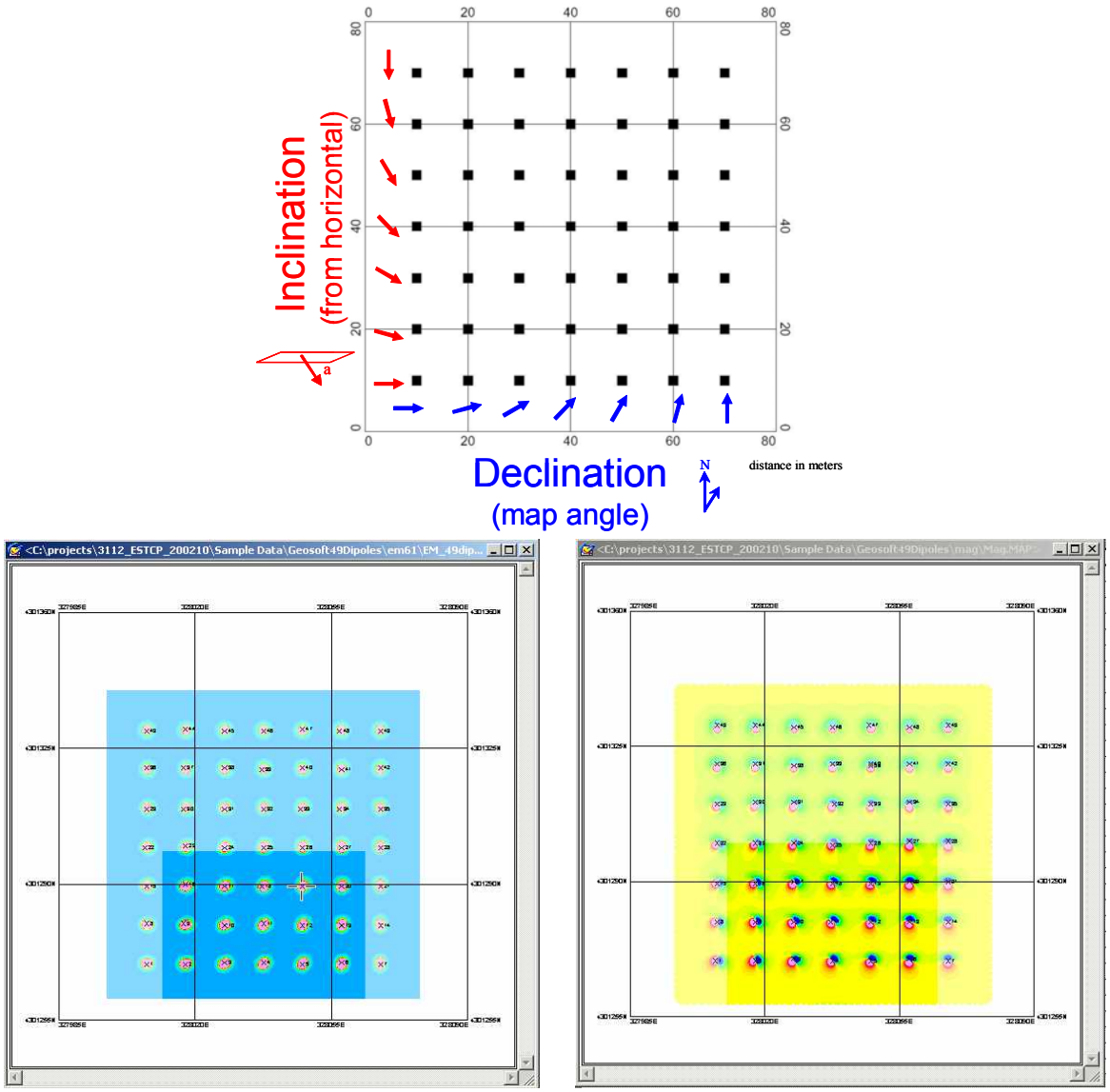


Figure 2-7. Top – schematic showing the inclination and declination of the source objects. The inclination and declination range from 0 to 90 degrees in 15-degree increments. Bottom – synthesized electromagnetic and magnetic data, left and right respectively.

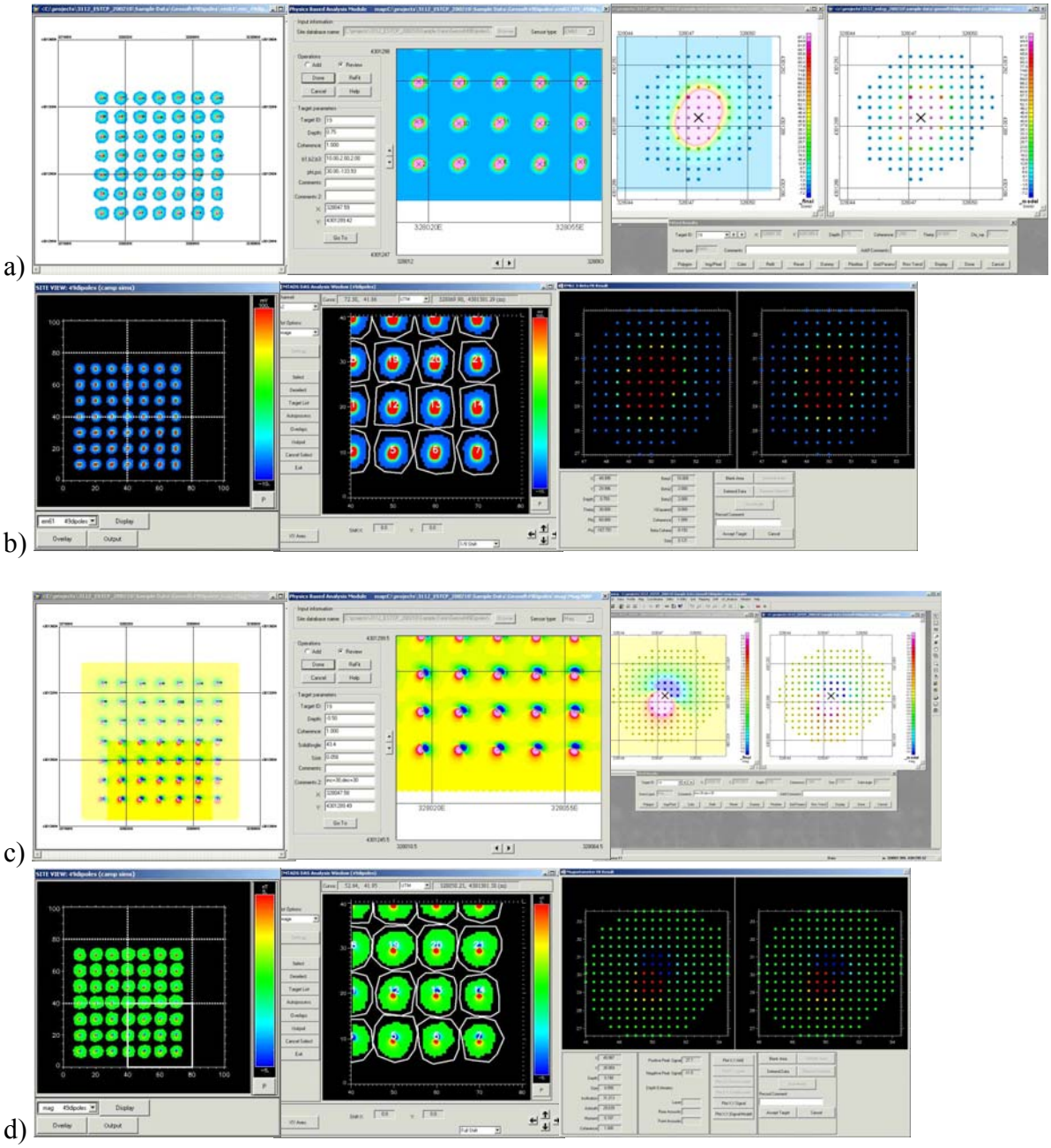


Figure 2-8. Screen snapshots showing from left to right along each row (i) the segmentation of individual anomalies, (ii) the graphical interface used to identify and select individual anomalies, and (iii) anomaly-specific comparisons of measured and modeled data. Row (a) shows images for UX-Analyze EMI data, (b) DAS EMI, (c) UX-Analyze magnetic, and (d) DAS magnetic.

2.2.3 Classification Modules included in this Demonstration

2.2.3.1 Classifiers

KNN: The k-nearest neighbor (KNN) classification rule uses the k neighbors of an unlabeled test point to estimate its label. The posterior probability of the class H given the unlabeled test point x, $p(H|x)$, is approximated by the proportion of the k neighbors in the labeled training data that are from class H. Thus, KNN provides a simple and intuitive classification rule where new data points are labeled according to a majority-vote of the nearest neighbors. KNN has been shown to have attractive asymptotic properties when the amount of available training data and the number of neighbors k are both large.

GLRT: The generalized likelihood ratio test is grounded in Bayesian decision theory. The likelihood ratio test $\lambda(x)$ is defined at the ratio of two conditional probabilities: the probability of the features given H1 $p(x|H1)$ and the probability of the features given H0 $p(x|H0)$, where H1 and H0 correspond to the UXO and clutter classes, respectively. Given probability distributions of the feature values for the classes of UXO and clutter, the likelihood ratio test can be calculated for any new data point by taking the ratio of the likelihoods of the features of the new data point under both H1 and H0. However, one issue arising with the likelihood ratio test is that the probability distributions $p(x|H1)$ and $p(x|H0)$ are rarely known. Thus, one solution is to assume the conditional probability distributions follow a parametric form with parameters estimated using the training data. The GLRT is calculated using the conditional probabilities that are dependent on the estimated parameters (McDonough and Whalen, 1995). In our implementation it is assumed that the UXO features have a Gaussian distribution.

SVM and RVM: The support vector machine (SVM) and relevance vector machine (RVM) are generalized linear classifiers. In both the SVM and RVM, the use of a kernel to represent the input data introduces nonlinearity and can transform the data into a higher dimensional space where it may be separable by a hyperplane. The differences between the SVM and RVM arise in the frameworks for optimization and training. The RVM finds relevance vectors that typically are located near the "centroids" of the decision boundary contours, whereas the SVM finds the support vectors that define the decision boundary. The RVM also does not utilize a margin between the classes, which is directly optimized in the training of the SVM. The SVM finds a decision boundary with the constraint of maximizing the margin, whereas the RVM does not consider a boundary margin in any sense. Instead, the RVM is a Bayesian kernel machine that applies a Bayesian framework to define the weights and relevance vectors through iterative calculation of the posterior weight distributions. A characteristic common to both the RVM and SVM is sparse representation of the decision space using a small subset of the training data. Rather than keeping track of all of the training data, the RVM and SVM techniques only require a limited subset of training vectors to discriminate between classes. The RVM tends to select fewer relevance vectors than the number of support vectors found by the SVM. Therefore, the training data is represented by an even more compact set of vectors which can further reduce the risk of overtraining.

2.2.3.2 Feature normalization

The feature normalization consisted of three steps. The feature values were first log-scaled (to prevent taking the log of negative values, the feature values were shifted to have a minimum value of 1 prior to taking the log). The feature values were normalized by the standard deviation calculated over the interquartile values (25 percentile to 75 percentile). We subtract the median value as the final step. Using the standard deviation of the interquartile range, rather than the standard deviation of all values, and subtracting off the median rather than the mean, makes the normalization more robust in the presence of outliers (i.e. outliers don't skew the normalized features). Each feature was normalized independently.

2.2.3.3 Feature selection

Feature selection was implemented using wrapper techniques, measuring classification performance on a set of training data with various sets of features to find the best-performing feature set. The feature selection technique was performed using four classifiers described above. We performed an exhaustive feature search for the best feature set containing between 1 and 5 features. The wrapper techniques used k-folds cross-validation with $k = 5$. The fitness metric for each feature set or feature, used to determine the best feature set in the exhaustive search or which feature to add next in the search, was the area under the ROC curve. If two feature sets / features have the same area under the ROC curve, the probability of false alarm at a probability of detection equaling 90% was used as a tie-breaker. The number of features was selected to maximize the fitness metric (area under the ROC curve) on the training data.

2.2.3.4 UX-Analyze classification capability

Dr. Leslie Collins, Duke University, and her colleagues incorporated the GLRT classifier into UX-Analyze as part of this program. The graphical user interface is shown in Figure 2-9.

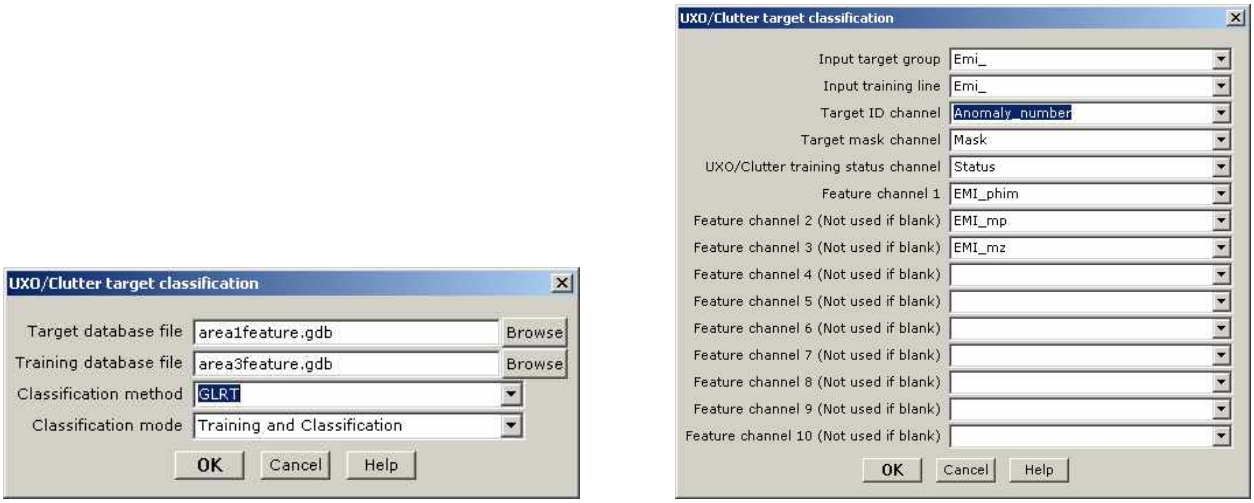


Figure 2-9. Screen snapshots showing the GLRT classification dialogue in UX-Analyze.

2.2.4 Data Analysis Documentation

UX-Analyze produces individualized anomaly reports, one for each anomaly, to document the decision process for each anomaly (Figure 2-10). In each plot, the measured data is graphically displayed next to the modeled data. The model parameters are listed in the middle of each page, and a profile extracted along the transect that passes closest to the dipoles location – as estimated by the inversion routine – is located at the bottom. The positions of individual measurements are superimposed on the maps.

Essentially, the anomaly plots graphically provide an intuitive confidence measure. If the measured and modeled data are indistinguishable, the reviewer can have confidence that the estimated source parameters are approximately correct. If the two maps are do not resemble each other, however, it tells us that the source in question (i) cannot be represented well using a point dipole source, (ii) is not isolated, (iii) does not have sufficient signal-to-noise ratio, or (iv) was not properly sampled (spatially or temporally). In any case, if the two maps are dissimilar the inverted model parameters are most likely not correct.

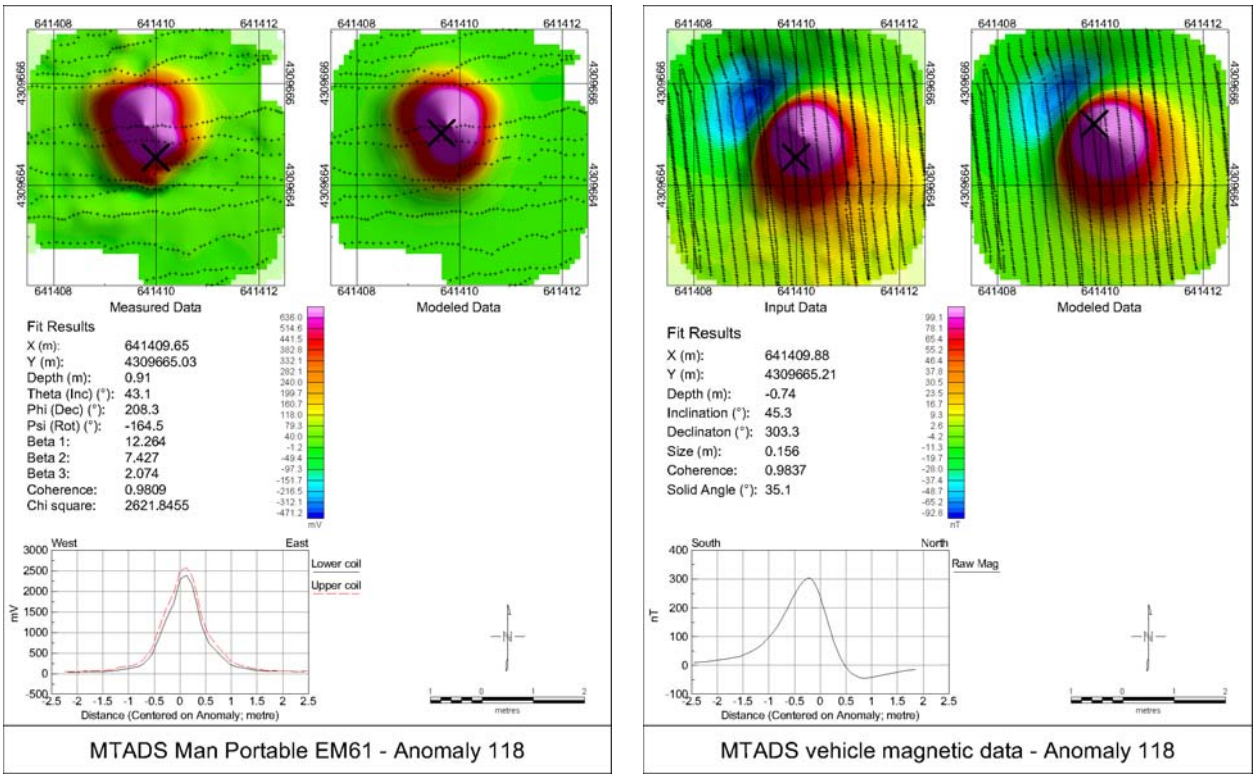


Figure 2-10. UX-Analyze generates a one-page summary for each anomaly. In the anomaly summaries shown above, the measured data is shown in the upper left hand corner, the inverted model parameters in the middle left, the forward model in the upper rights, and a profile in the lower left corner. EMI data for the anomaly are shown in the left summary, and magnetic data on the right.

2.3 Factors Affecting Cost and Performance

The analysis approach demonstrated here utilizes the spatial distribution of the measured magnetic or EMI signatures. As such, it requires high signal-to-noise data that possess a high degree of spatial precision across the footprint of the anomaly. The costs to acquire data that will support discrimination decisions are higher than that required if the goal is only to detect the presence of an object. The analysis costs are also higher if attempts are made to quantitatively discriminate relative than only to detect.

The factors affecting acquisition costs relate to particulars of the sensing system, spatial registration system, the target objectives, and the site environment. Although these costs are not the focus of this demonstration, they are very important to the ultimate transferability of this approach.

The factors affecting analysis time include are significantly affected by (i) the degree to which the anomalies are spatially separated, (ii) the number of anomalies, and (iii) the amount of geologic-

related signatures that possess similar wavelengths as the targeted signatures. The data density is also a factor, but only marginally so compared to the factors listed above.

Discrimination performance is measured by our ability to characterize and classify one object from another. The factors that affect performance, therefore, relate to the similarity (in feature space) between the sought-after object versus the clutter, our ability to accurately measure the responses, the presence of signatures that spatially interfere or otherwise compete with the UXOs response, as well as our ability to quantitatively characterize and classify the source objects. Many of these factors are not under our direct control.

2.4 Advantages and Limitations of the Technology

This analysis approach uses spatially referenced geophysical data to estimate target parameters for each anomaly. This has an inherent advantage over ancillary analysis methods that are commonly used. Due to a lack of analysis routines available, many contractors make UXO and non-UXO declarations based on anomaly amplitude, half width, spatial footprint, or overall 'look'. These characterization methods are sensitive to the targets' orientation and depth of burial (or distance from the sensor). The methodology demonstrated here separates the measured signatures into that which is inherent to the target, and that which is related to the geometry of the problem (such as distance to sensor and orientation). The fitted parameters that are inherent to the target itself are used to classify the unknown object.

The primary advantage, therefore, is the potential for discriminating between UXO and non UXO-like objects based upon geophysical survey data. This is in contrast to simply identifying the location of anomalies from the geophysical survey data. Magnetic discrimination is based primarily on the apparent fitted dipole size (or scaled dipole moment). Using EMI data, increased discrimination performance can sometimes be achieved by utilizing estimated shape information. If successful discrimination capabilities can be achieved, significant excavation savings can be realized by leaving the non-hazard clutter items unearthed.

This is not to say, however, that the data analysis technology being demonstrated will solve the UXO characterization and classification problems. Even with optimal data quality, the estimated fit parameters cannot always be separated into distinct, non-overlapping classes of UXO and non UXO-like objects. In fact, none of the fit parameters are actually unique to UXO items. Because of this, clutter items that physically resemble UXO will probably be misclassified. Additionally, if the data quality is not optimal the fitted parameters cannot be trusted.

3 Demonstration Design

3.1 Performance Objectives

Table 3-1. Qualitative Performance Objectives

Type of Performance Objective	Primary Performance Criteria	Expected Performance (metric)	Actual Performance Objective Met?
Qualitative	Ease of use	Minimal training required for data processor experienced in Oasis montaj	Minimal training is required for users experienced in Oasis montaj
	Robustness	Analysis flow not seriously interrupted by bugs	Software bugs initially limited/stopped flow, but were fixed during the course of the demonstration
	End-user Acceptance	NA	Positive
Quantitative	Non-TOI Rejection Rate	0.25 or greater	Yes
	TOI Retention Rate	0.8 or greater	Yes
	Location Accuracy	0.3m	Yes
	Depth Accuracy	0.3m	Yes

3.2 Selecting Test Sites

The APG site was selected because of the combination of reliable ground truth information and availability of multiple, high quality production data sets.

3.3 Test Site History/Characteristics

The Standardized UXO Technology Demonstration Site Program is a multi-agency program spearheaded by the U.S. Army Environmental Center. The U.S. Army Aberdeen Test Center and the U.S. Army Corps of Engineers Engineer Research and Development Center provide programmatic support. The program was funded and supported by ESTCP, SERDP, and the Army Environmental Quality Technology program. A brief description of the APG test site is presented below. See http://aec.army.mil/usaec/technology/uxo03a.html for additional details.

The APG Standardized Test Site is located within a secured range area of the Aberdeen Proving Ground, which is located approximately 30 miles northeast of Baltimore at the northern end of the Chesapeake Bay. The Standardized Test Site encompasses 17 acres of upland and lowland flats, woods and wetlands. There are multiple challenge areas within APG. The blind grid and open field areas are shown in Figure 3-1.

The Open Field area provides the demonstrator with a variety of realistic scenarios essential for evaluating sensor system performance. The scenarios and challenges found on the Open Field area

consist of a gravel road, wet areas, dips ruts and trees. Vegetation height varies from 15 to 25 centimeters. Other challenges within the open field site include electrical lines, swales, stone pads/roads, and metallic fencing. The Standardized UXO Technology Demonstration Site Program emplaced a diverse set of targets at the APG test site (Table 3-2).

Table 3-2. Inert ordnance emplaced at APG.

Standard Type	Nonstandard (NS)
20-mm Projectile M55	20-mm Projectile M55
	20-mm Projectile M97
40-mm Grenades M385	40-mm Grenades M385
40-mm Projectile MKII Bodies	40-mm Projectile M813
BDU-28 Submunition	
BLU-26 Submunition	
M42 Submunition	
57-mm Projectile APC M86	
60-mm Mortar M49A3	60-mm Mortar (JPG)
	60-mm Mortar M49
2.75-inch Rocket M230	2.75-inch Rocket M230
	2.75-inch Rocket XM229
MK 118 ROCKEYE	
81-mm Mortar M374	81-mm Mortar (JPG)
	81-mm Mortar M374
105-mm Heat Rounds M456	
105-mm Projectile M60	105-mm Projectile M60
155-mm Projectile M483A1	155-mm Projectile M483A
	500-lb Bomb

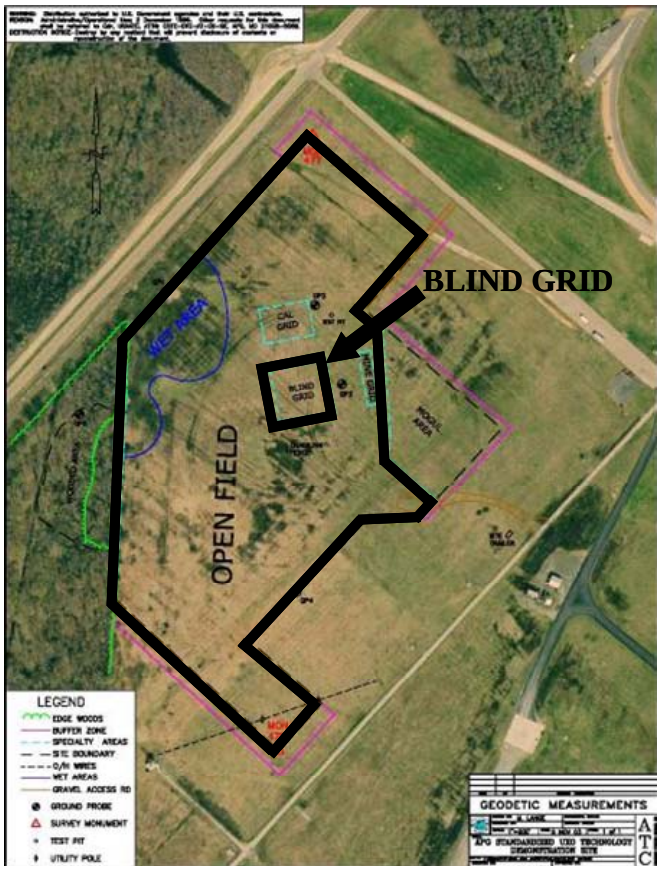


Figure 3-1. Aerial photograph of the Aberdeen Proving Ground Test Site.

3.4 Present Operations

The Standardized Test Sites have been, and continue to be, utilized to benchmark a significant number of technologies and contractors (McDonnel and Karwatka, 2007).

3.5 Pre-Demonstration Testing and Analysis

This demonstration used characterization and classification algorithms that have been prototyped and undergone limited testing during previous research programs. Characterization algorithms for magnetic and electromagnetic (EM61 MK1 and MK2) and classification routines were rewritten and compiled into dynamic link libraries and integrated into Oasis montaj. Prior to the demonstration, we verified that the rewritten code produces the same result as the prototype code for each algorithm.

3.6 Testing and Evaluation Plan

3.6.1 Demonstration Set-up and Start-up

This demonstration leveraged data acquired under ancillary projects. Details of the hardware setup and field operations can be found at http://aec.army.mil/usaec/technology/uxo03f01.html. The relevant reports are SR0157 TTFW and SR0311 G-tek.

3.6.2 Period of Operation

Data collection occurred under prior projects associated with the Standardized Test Site program. TTFW acquired EM61 data in November 2003. Geo-Centers acquired magnetic and EMI data in August 2004. G-tek acquired magnetic data in October 2003. NRL acquired magnetic data during June 2004 and EMI data during October 2004.

3.6.3 Area characterized or Remediated

The APG Standardized UXO Technology Demonstration Open Field area comprises 13.7 acres and the Blind Grid comprises 0.5 acres. No targets were remediated as part of this demonstration.

Two collections, or groups, of data from APG are considered here. The first collection consists of stationary data that were acquired using a template grid over 250 anomalies (identified in Figure 3-2) from the Opend Field and 20 anomalies from the Blind Grid/Calibration area. The selected targets were also surveyed by TTFW and NRL's EMI systems. The stationary data provide a baseline, therefore, with which to compare results obtained from the dynamic production data using common targets.

The second target collection consists of anomalies in the open field (Figure 3-3) that were selected, analyzed, and scored in a blind test. Magnetic and EMI analysis results for the second target selection were sent to the AEC and IDA for scoring. The template data, and their labels, were acquired after blind tests were completed and scored. They were not used in any way to better the performance of the blind test.

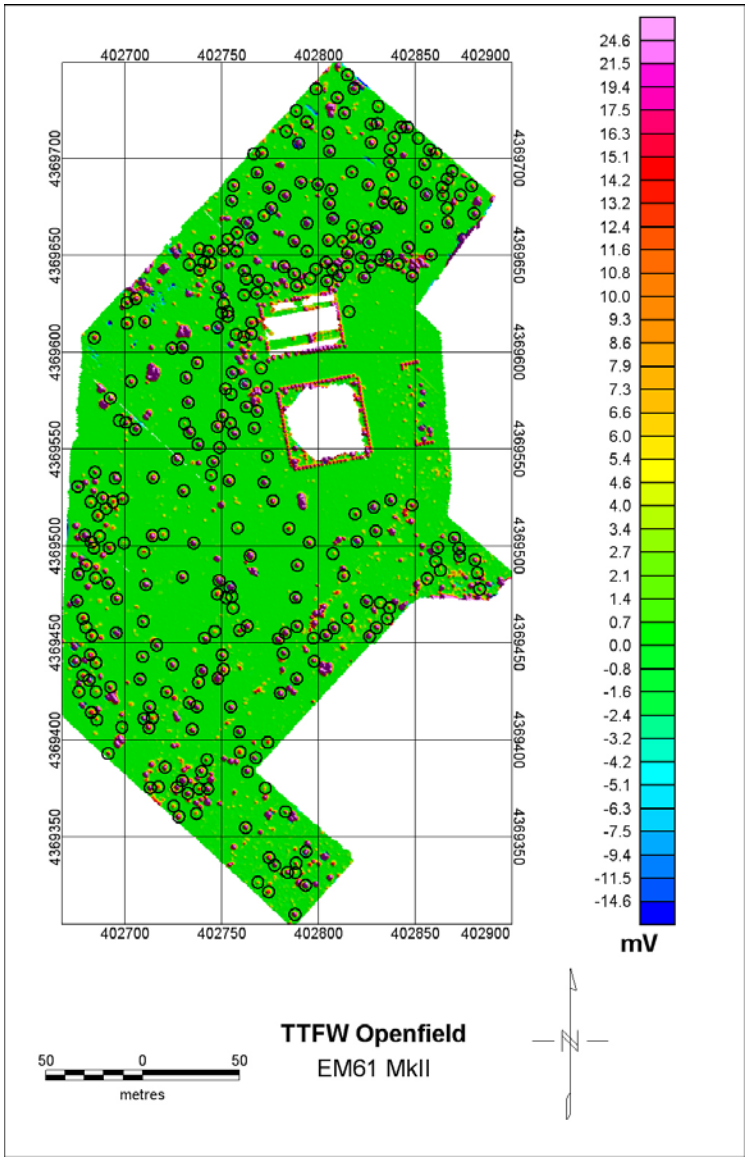


Figure 3-2. Color contour map of the EM61 Mk2 data (0.366ms time gate) collected by TTFW in November 2003. The symbols represent targets for cued investigation.

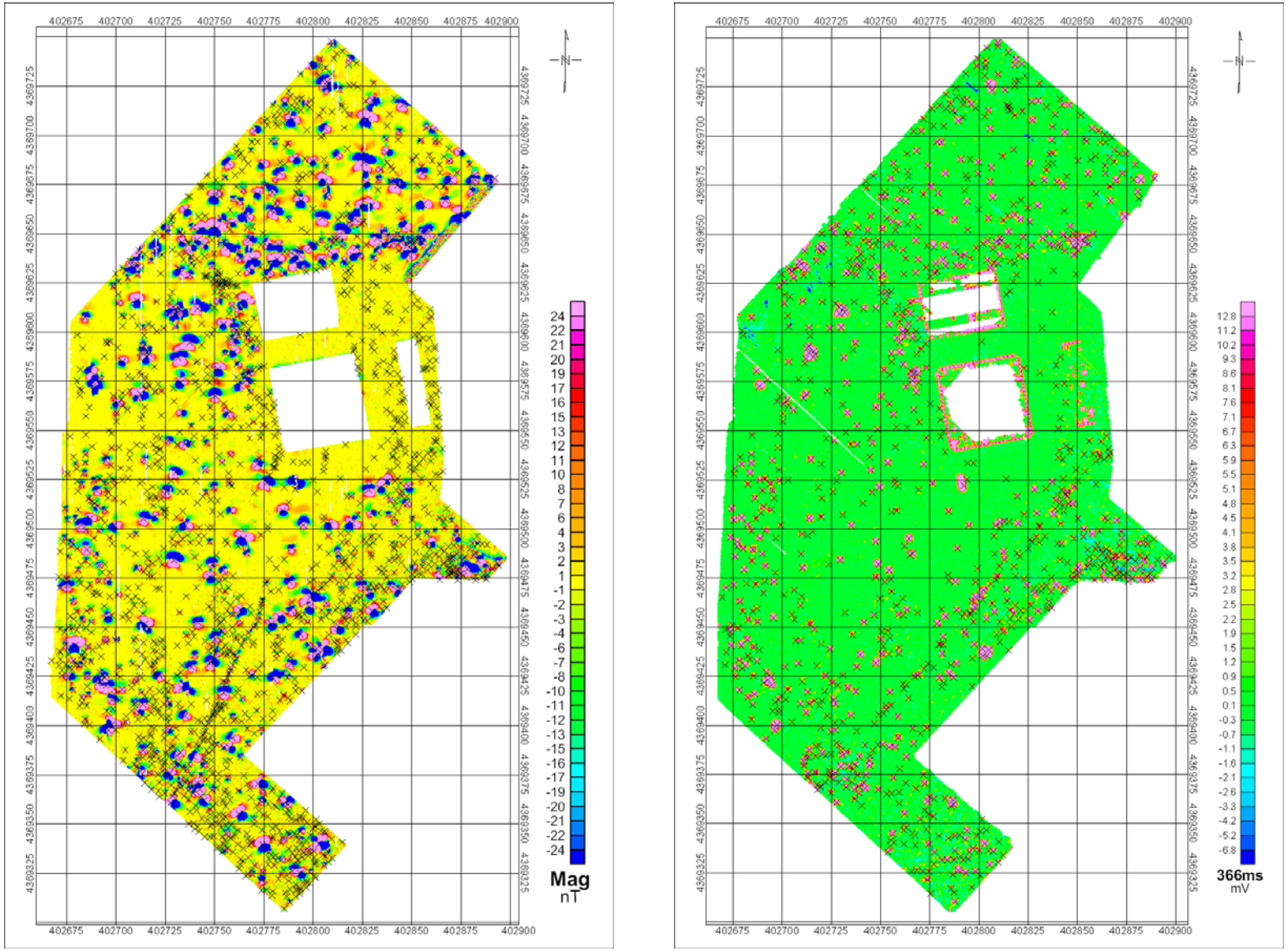


Figure 3-3. False color images showing G-tek magnetic (left) and TTFW EMI (right) data from the Open Field area. The symbols identify anomalies included in the analysis.

4 Performance Assessment

4.1 Performance Criteria

Table 4-1. Criteria for this Demonstration

Performance Criterion	Description	Primary or Secondary
Ease of Use	Analysis flow and anticipated skill level required.	Primary
Robustness	No major bugs that artificially limit the analysts' ability to conduct analysis.	Primary
End-user Acceptance	The EPA and Corps of Engineers must ultimately accept and prefer this approach if it is going to be successfully transitioned. This will be our first opportunity for constructive criticism.	Primary
Non-TOI Rejection Rate	Number of non-TOIs rejected by the discrimination decision / Number of non-TOIs	Primary
TOI Retention Rate (Efficiency)	Number of TOIs retained after discrimination / Number of TOIs.	Primary
Location Accuracy	The average absolute distance error for ordnance correctly identified in the discrimination stage	Primary
Depth Accuracy	The average missed depth for ordnance correctly identified in the discrimination stage	Primary

The 'ease of use' and 'robustness' criterion report the analysts' experience regarding the programs' flow, speed, computational capabilities, graphics, and bugs as they relate to efficient data analysis.

Analysis time was logged manually. The principal baseline metrics are (i) identifying, characterizing, and classifying anomalies, and (ii) documenting the analysis with anomaly specific fit images, parameters, and dig sheet information. Actual times for these metrics are reported.

4.2 Performance Confirmation Methods

Table 4-2. Quantitative Performance Objectives: Open Field Blind Test (MAG GLRT)

Primary Performance Criteria	Expected Performance Metric (pre demo)	Performance Confirmation	Actual Performance Open Field
PRIMARY CRITERIA (Performance Objectives) (Quantitative)			
Non-TOI Rejection Rate	0.25 or greater	Open Field	0.27
TOI Retention Rate	0.8 or greater	Open Field	0.87
Location Accuracy	0.3m	Open Field	0.24m
Depth Accuracy	0.3m	Open Field	0.18m

Table 4-3. Quantitative Performance Objectives: Open Field Blind Test (EMI GLRT)

Primary Performance Criteria	Expected Performance Metric (pre demo)	Performance Confirmation	Actual Performance Open Field
PRIMARY CRITERIA (Performance Objectives) (Quantitative)			
Non-TOI Rejection Rate	0.25 or greater	Open Field	0.25
TOI Retention Rate	0.8 or greater	Open Field	0.82
Location Accuracy	0.3m	Open Field	0.24m
Depth Accuracy	0.3m	Open Field	0.01m

4.3 Data Analysis, Characterization, and Evaluation

4.3.1 Cued Target Collection Comparisons and Results

Approximately 250 targets from the APG test site were surveyed using both a stationary template and dynamic survey modes. We received labels for 107 of these. We consider here two dynamic surveys, the first collected using an array of EM61 sensors, and the second used a commercial EM61 cart system. In this section, we present visualizations of the fitted parameters and present ROC curves for labeled targets. This comparison and analysis of different data sets, with different inherent data quality, provides a baseline and establishes performance bounds.

In Figure 4-1, we plot SNR versus dipole fit error for all of the targets that were surveyed using the template. Symbols and colors identify the data type. Black diamond symbols are used for the template, blue asterisks identify EM61 Array data, and red triangles are used to show EM61 Cart data. As expected, the stationary template data possess much smaller dipole fit errors [defined as 100*sqrt(1-correlation_coefficient^2); Table 4-4].

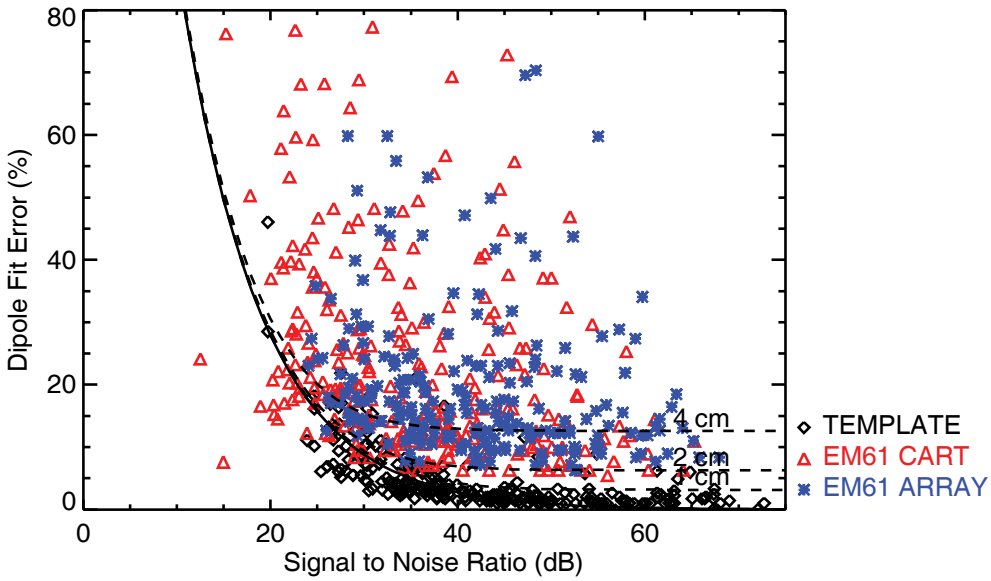


Figure 4-1. SNR versus dipole fit errors for all targets surveyed using the template.

Table 4-4. SNR and Fit Error statistics across data sets

	SNR Mean	SNR Standard Deviation	Dipole Fit Error Mean	Dipole Fit Error Standard Deviation
Template (stationary)	42.9	12.3	4.7	5.1
TTFW (dynamic)	35.7	10.8	24.1	17.2
NRL (dynamic)	41.6	10.7	19.4	11.9

Actual versus estimated depths for the cued target collection anomalies are shown in Figure 4-2. Depth errors for UXO only are shown in Figure 4-3.

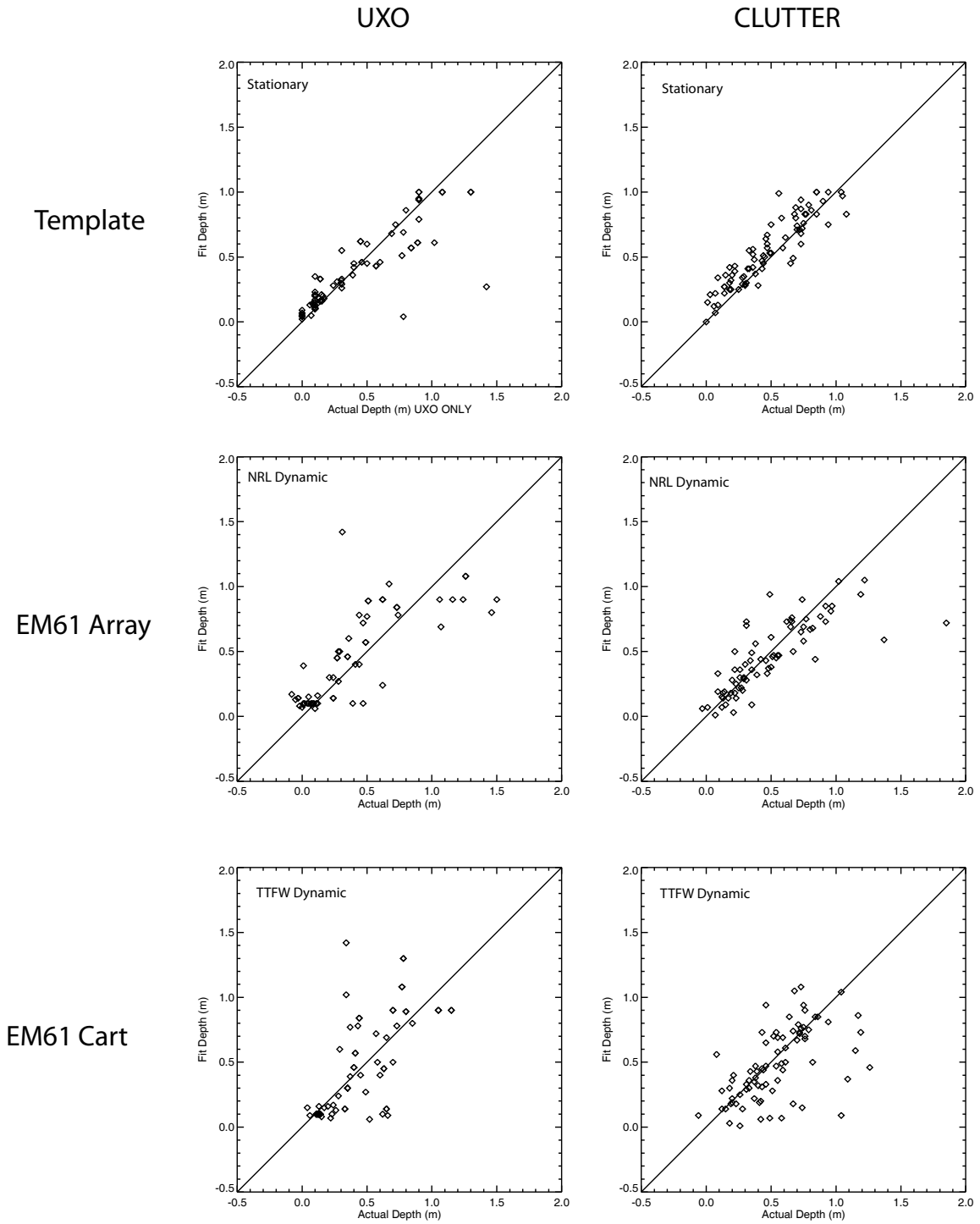


Figure 4-2. Actual versus fitted depth estimates.

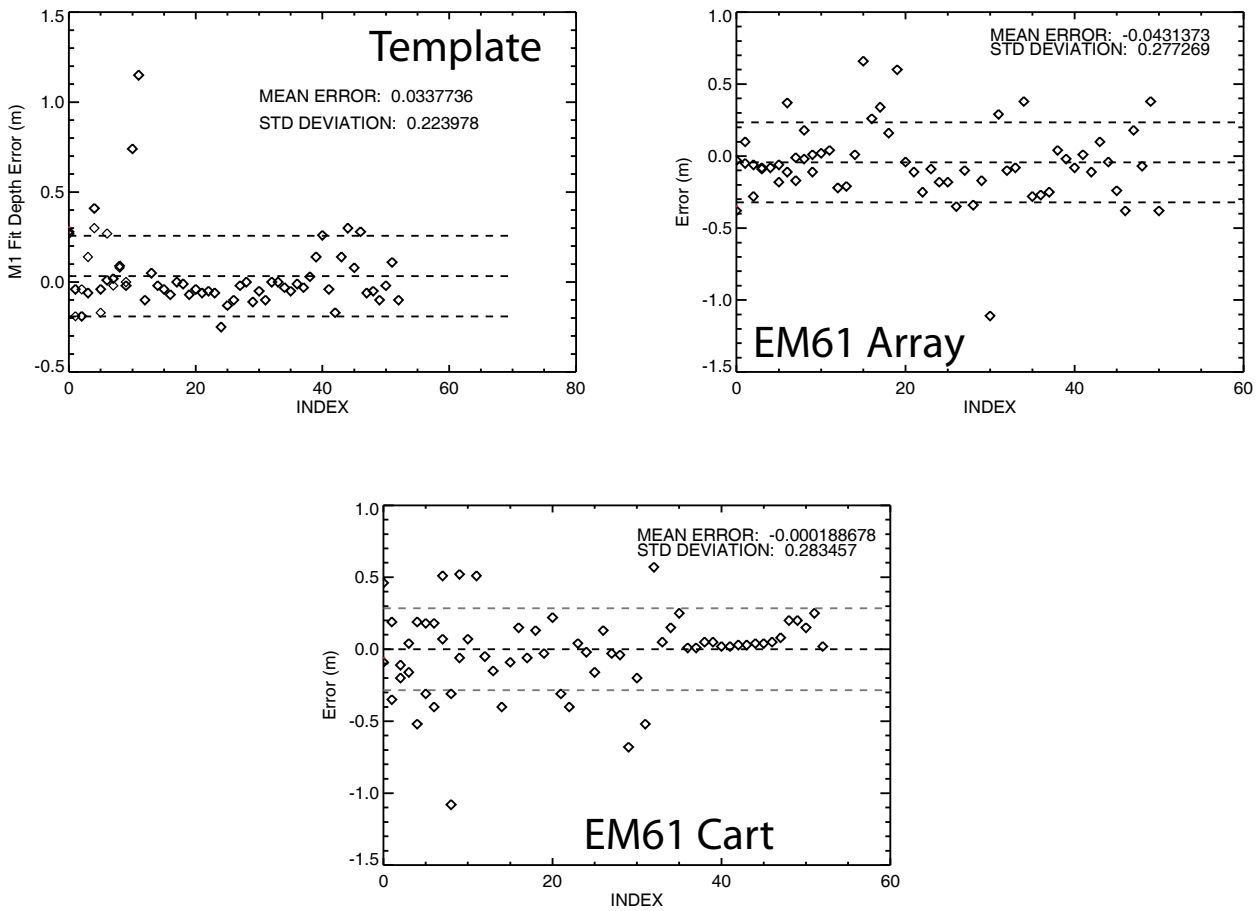


Figure 4-3. Depth errors for UXO

Polarizations for UXO are shown in Figure 4-4. The vertical lines in the figure extend from β2 and run to β3 for each fitted target. Because the UXO are axially symmetric, we expect that β1>β2≈β3. This is generally true for the template data, but not for the cart mounted EM61 and marginally true for the EM61 array data.

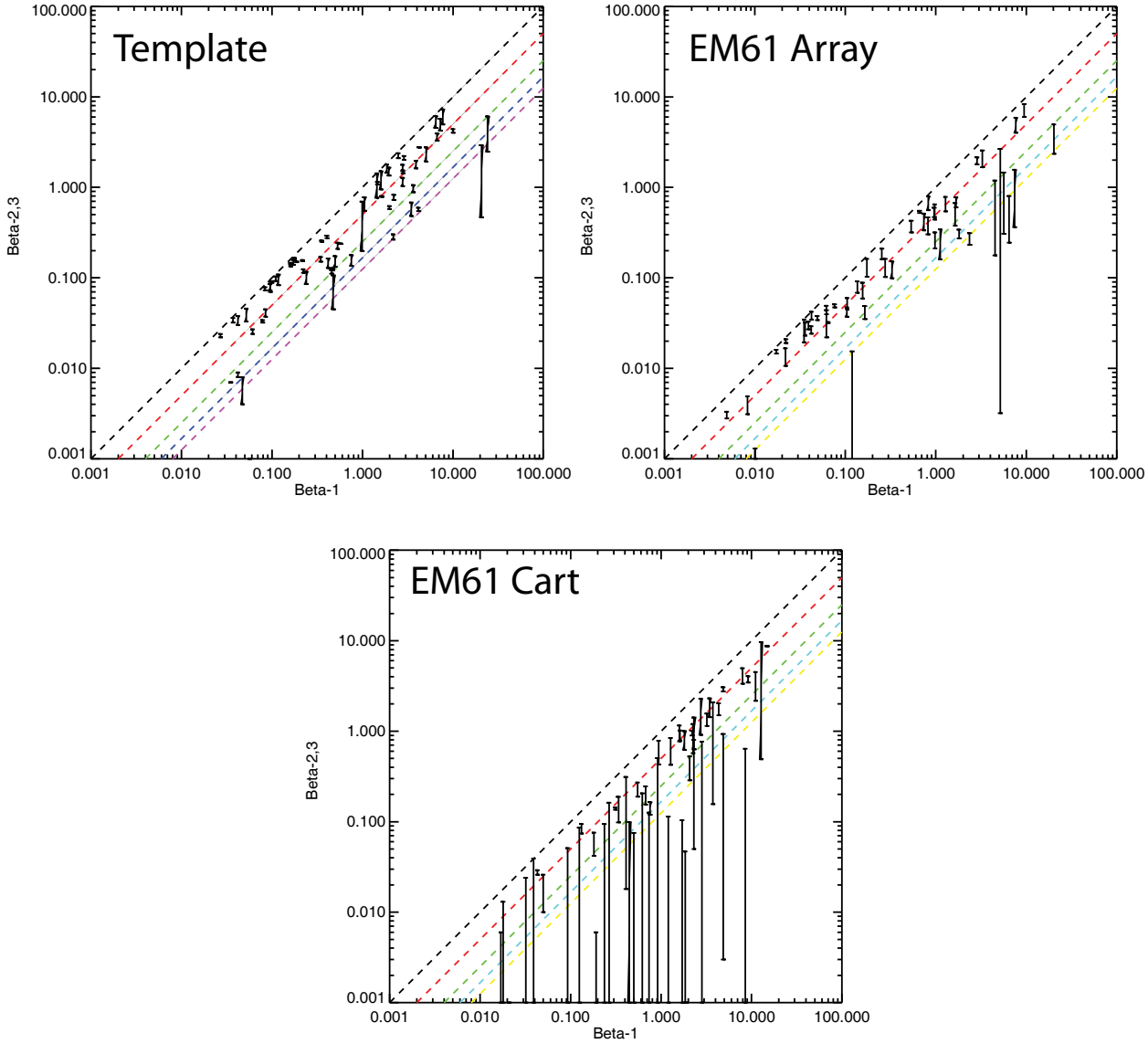
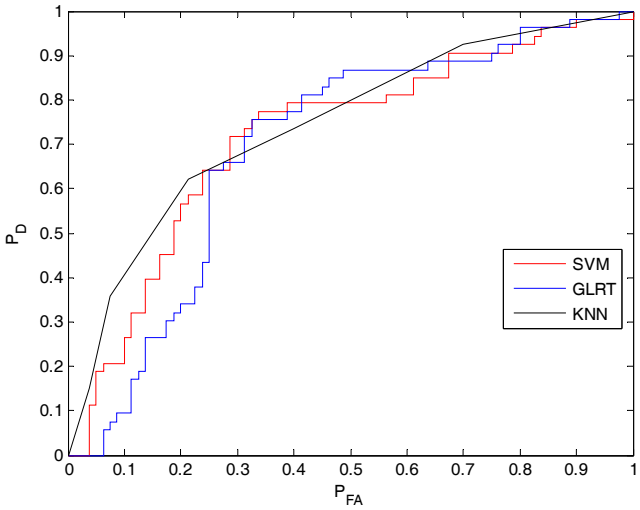


Figure 4-4. Polarization plots for labeled UXO from the Open Field.

ROC curves for the cued target collection are shown in Figure 4-5 thru Figure 4-7. For each classifier, we considered intrinsic features only. An exhaustive search of features was then performed, considering (2^N-1) sets of features, where N is the number of intrinsic features. We selected the set of features that maximized area under the ROC curve. In the figures, the accompanying table shows the area under the curve if the features selected with one classifier (rows) are used on a different classifier (columns).

Based on the area under the curve (AUC) for the ROCs, the template data performed best, followed by the EM61 array results. The EM61 Cart data performed the poorest. Results of the template data, however, were not ideal. Although the majority of the UXO were classified as being UXO, some of larger UXO were classified as being very similar to non-TOIs. Inspection into the template-data failures showed that the 9 out of the first 10 were either 105mm or 155mm projectiles.

The reason for misclassifying large UXO is perhaps explained by examining the distribution of clutter at this site. Figure 4-8 compares the polarization estimates by data set and by UXO size. In the figure, black color is used to identify clutter (non-TOI) and red color is used for UXO. Also, 'Small UXO' refers to ordnance smaller than 60mm in diameter; 'Large UXO' refers to 105mm UXO and larger; and 'Medium UXO' is everything else. From this plot, it is clear the emplaced clutter at APG significantly overlaps, in polarization feature space, with large UXO items. The clutter marginally overlaps with the medium-sized UXO, and not all with small-size UXO. ROC curves for the three size-based clusters of UXO are presented in Figure 4-9 thru Figure 4-11.



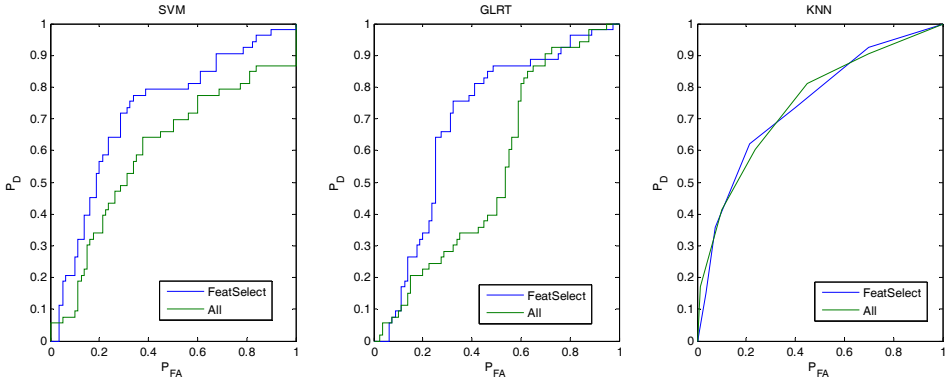
Classifier for ROC

Classifier for feature selection

	SVM	GLRT	KNN
SVM	0.718632	0.701651	0.730778
GLRT	0.659198	0.694104	0.711085
KNN	0.60684	0.514387	0.744929

Area under the ROC curve

a)



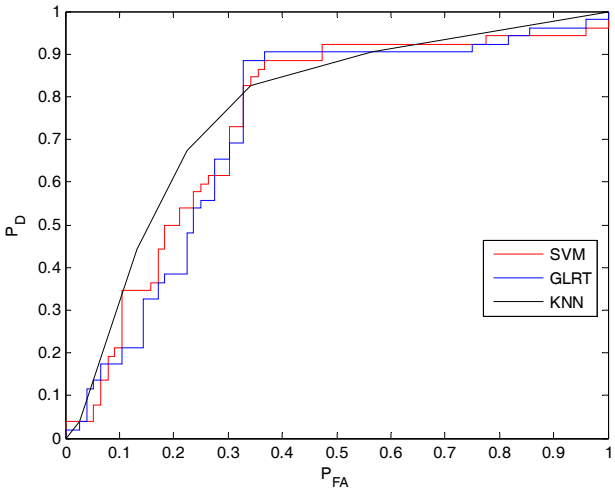
b)

SVM	GLRT	KNN	All intrinsic
TTFW Fit_b1	TTFW Fit_b2	TTFW Fit_b1	TTFW Fit_b1
TTFW Fit_b2	TTFW Fit_b3	TTFW Fit_b3	TTFW Fit_b2
TTFW Fit_b1_TTFW Fit_b2_TTFW Fit_b3_sum		TTFW Fit_b1_TTFW Fit_b2_R	TTFW Fit_b3
		TTFW Fit_b2_TTFW Fit_b3_R	TTFW Fit_b1_TTFW Fit_b2_R
			TTFW Fit_b2_TTFW Fit_b3_R
			TTFW Fit_b1_TTFW Fit_b2_TTFW Fit_b3_sum

c)

Figure 4-5. Discrimination performance for EM61 Cart data (cued target collection). A) ROC curve B) Comparison of all intrinsic features versus selected C) table listing selected and intrinsic features (names ending in "_R" or "_sum" are a ratio or sum respectively)

a)

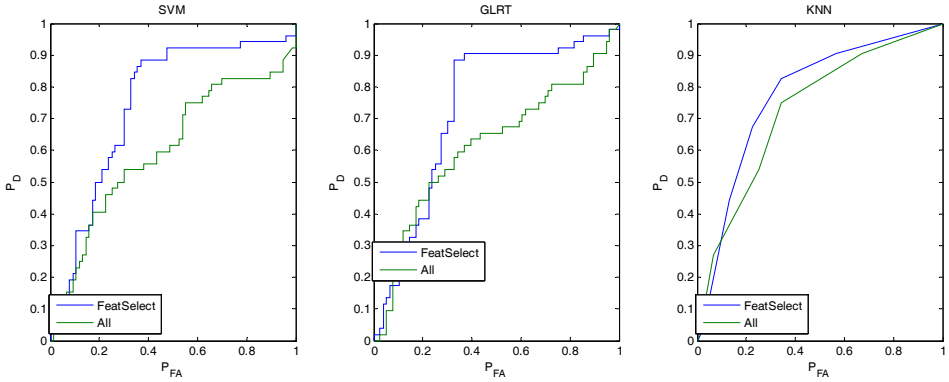


Classifier for ROC

Classifier for feature selection	SVM	GLRT	KNN
SVM	0.742662	0.726974	0.7729
GLRT	0.742409	0.729251	0.7729
KNN	0.742409	0.726721	0.7729

Area under the ROC curve

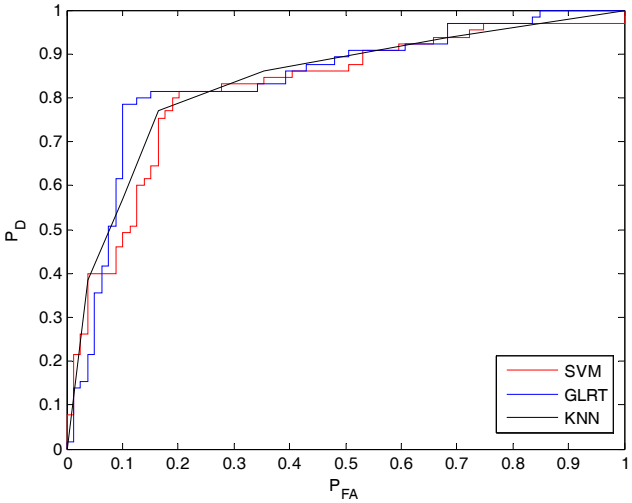
b)



c)

SVM	GLRT	KNN	All intrinsic
Fit_b3	Fit_b3	Fit_b3	Fit_b1
			Fit_b2
			Fit_b3
			Fit_b1_Fit_b2_R
			Fit_b2_Fit_b3_R
			Fit_b1_Fit_b2_Fit_b3_sum

Figure 4-6. Discrimination performance for EM61 Array data (cued target collection). A) ROC curve B) Comparison of all intrinsic features versus selected C) table listing selected and intrinsic features (names ending in "_R" or "_sum" are a ratio or sum respectively)



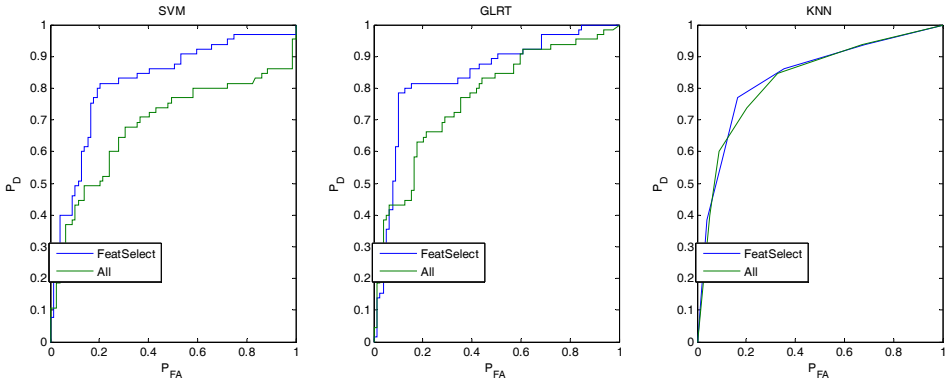
a)

Classifier for ROC

Classifier for feature selection

	SVM	GLRT	KNN
SVM	0.820837	0.822395	0.808179
GLRT	0.664654	0.840896	0.803116
KNN	0.776144	0.783642	0.837585

Area under the ROC curve



b)

SVM	GLRT	KNN	All intrinsic
Beta2	Beta2	Beta 2	Beta 1
Beta3	Beta3	Beta 3	Beta 2
BestCohRatio		BestCohRatio	Beta 3
		Beta2_Beta3_R	uncFitError
		Beta1_Beta2_Beta3_sum	BestCohRatio
			Beta1_Beta2_R
			Beta2_Beta3_R
			Beta1_Beta2_Beta3_sum

c)

Figure 4-7. Discrimination performance for EM61 Template data (cued target collection). A) ROC curve B) Comparison of all intrinsic features versus selected C) table listing selected and intrinsic features (names ending in "_R" or "_sum" are a ratio or sum respectively)

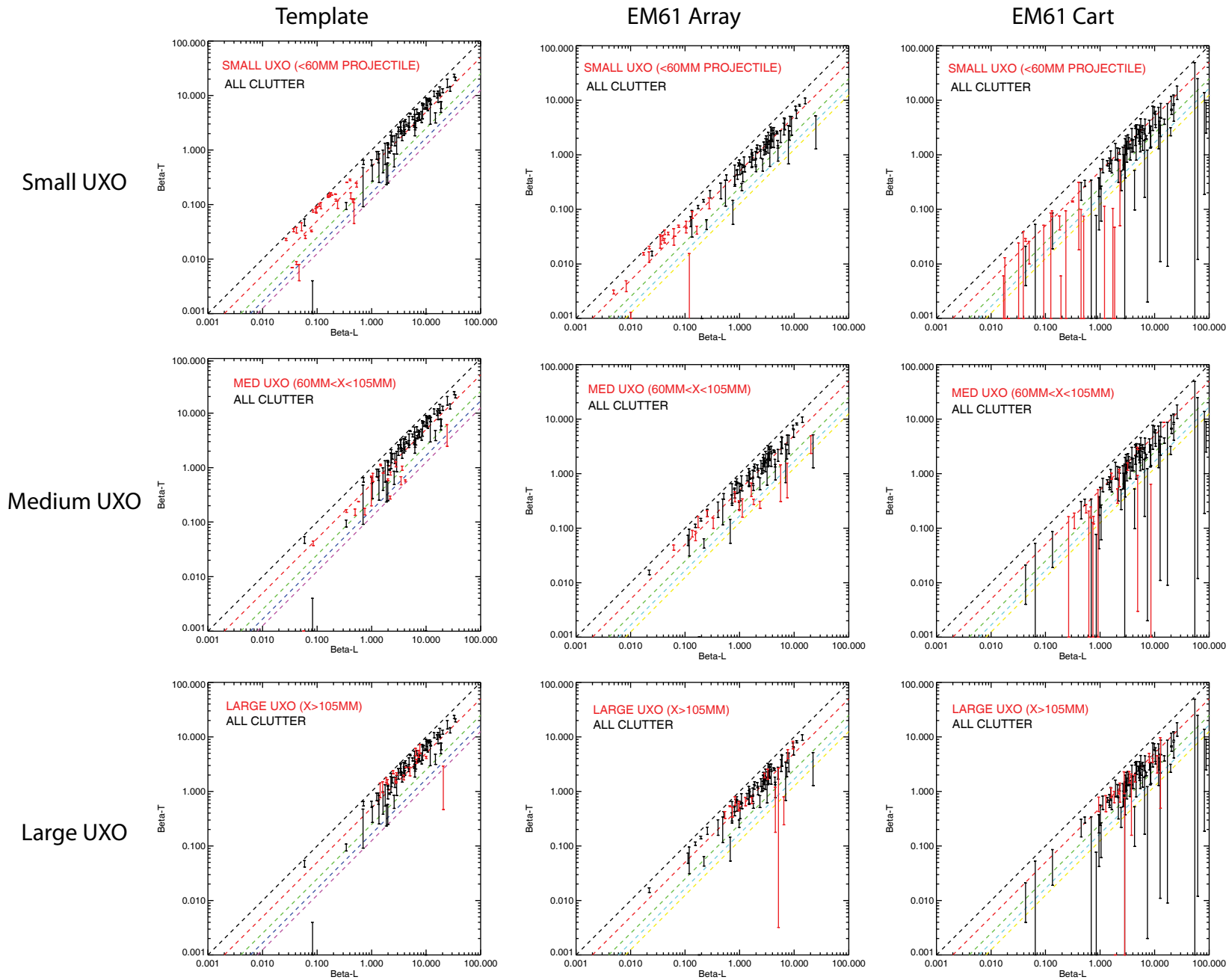
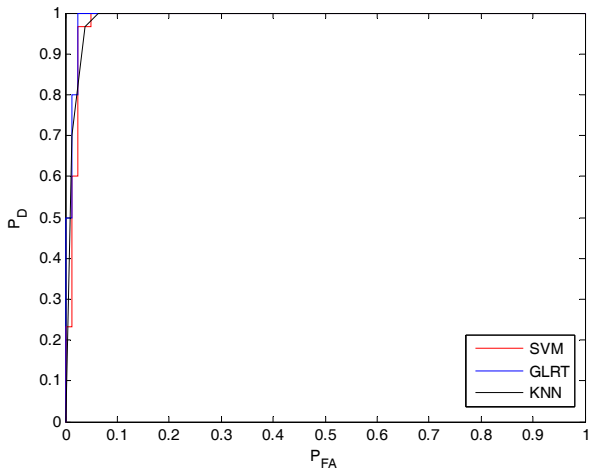


Figure 4-8. Polarization plots segmented by survey approach and by size-based clusters of UXO.

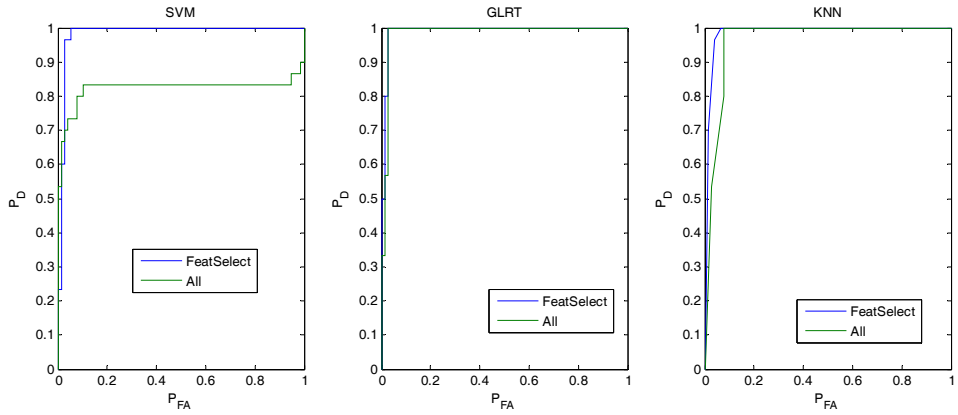


Classifier for ROC

Classifier for feature selection	SVM	GLRT	KNN
SVM	0.984388	0.987764	0.987131
GLRT	0.92616	0.991139	0.962236
KNN	0.980591	0.989873	0.987131

Area under the ROC curve

a)

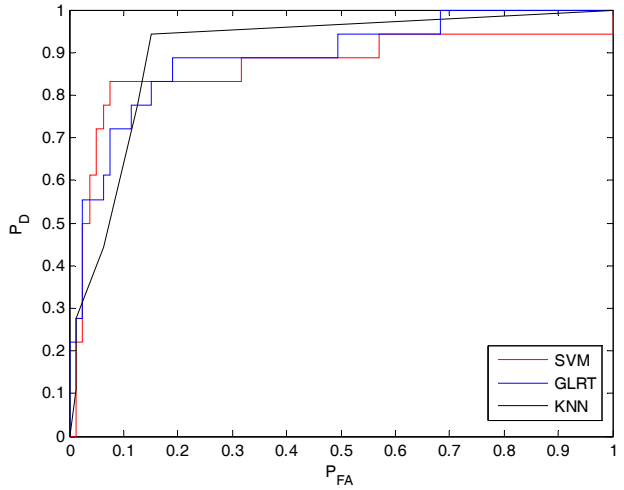


b)

SVM	GLRT	KNN	All intrinsic
Beta 1	Beta 2	Beta 1	Beta 1
Beta 2	uncFitError	Beta 2	Beta 2
uncFitError	Beta 2_Beta 3_R	uncFitError	Beta 3
Beta 1_Beta 2_Beta 3_sum	Beta 1_Beta 2_Beta 3_sum		uncFitError
			BestCohRatio
			Beta 1_Beta 2_R
			Beta 2_Beta 3_R
			Beta 1_Beta 2_Beta 3_sum

c)

Figure 4-9. Discrimination performance for EM61 Template data, restricting the TOI subclass to UXO smaller than 60mm. A) ROC curve B) Comparison of all intrinsic features versus selected C) table listing selected and intrinsic features (names ending in "_R" or "_sum" are a ratio or sum respectively).



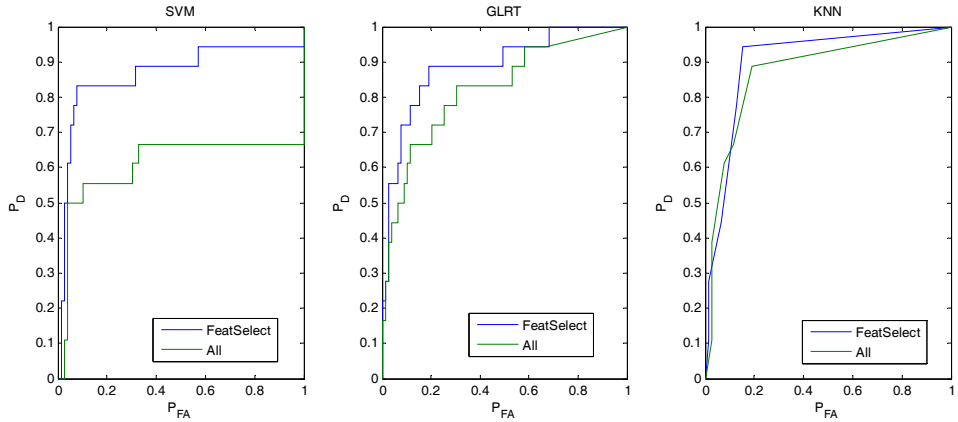
a)

Classifier for ROC

Classifier for feature selection

	SVM	GLRT	KNN
SVM	0.867792	0.869902	0.841772
GLRT	0.809423	0.889592	0.81294
KNN	0.618847	0.777778	0.904008

Area under the ROC curve

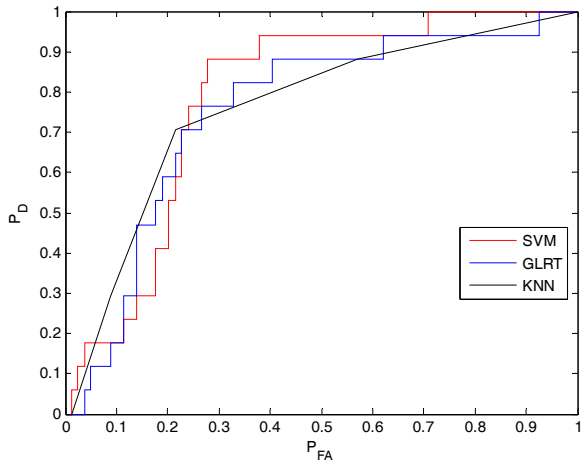


b)

SVM	GLRT	KNN	All intrinsic
Beta 2	Beta 1	Beta 1	Beta 1
Beta 3	Beta 3	uncFitError	Beta 2
uncFitError	uncFitError	Beta 1_Beta 2_R	Beta 3
	Beta 1_Beta 2_Beta 3_sum	Beta 1_Beta 2_Beta 3_sum	uncFitError
			BestCohRatio
			Beta 1_Beta 2_R
			Beta 2_Beta 3_R
			Beta 1_Beta 2_Beta 3_sum

c)

Figure 4-10 Discrimination performance for EM61 Template data, restricting the TOI subclass to UXO larger than 57mm but smaller than 105mm projectiles. A) ROC curve B) Comparison of all intrinsic features versus selected C) table listing selected and intrinsic features (names ending in "_R" or "_sum" are a ratio or sum respectively)



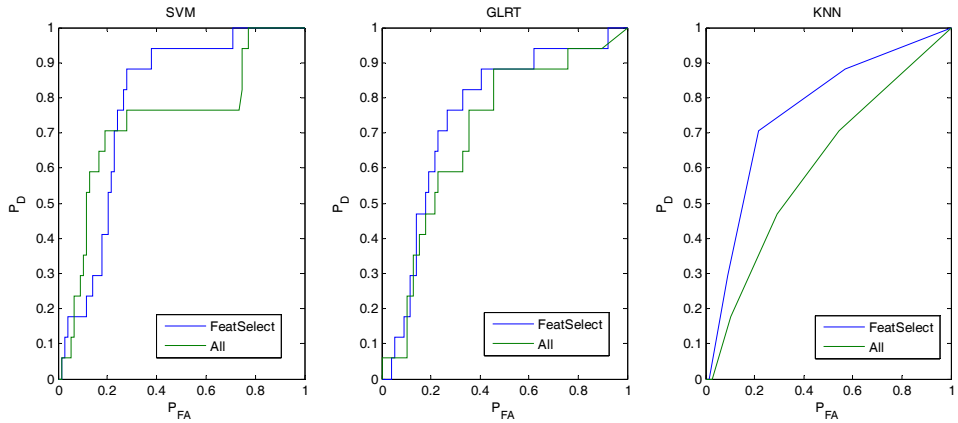
Classifier for ROC

Classifier for feature selection

	SVM	GLRT	KNN
SVM	0.786299	0.654505	0.609829
GLRT	0.714073	0.754281	0.548771
KNN	0.802308	0.643336	0.760983

Area under the ROC curve

a)



b)

SVM	GLRT	KNN	All intrinsic
Beta 2	Beta 3	Beta 1	Beta 1
uncFitError	uncFitError	uncFitError	Beta 2
Beta 2_Beta 3_R	Beta 1_Beta 2_R	BestCohRatio	Beta 3
Beta 1_Beta 2_Beta 3_sum	Beta 1_Beta 2_Beta 3_sum	Beta 2_Beta 3_R	uncFitError
Beta 2_Beta 3_R		Beta 1_Beta 2_Beta 3_sum	BestCohRatio
			Beta 1_Beta 2_R
			Beta 2_Beta 3_R
			Beta 1_Beta 2_Beta 3_sum

c)

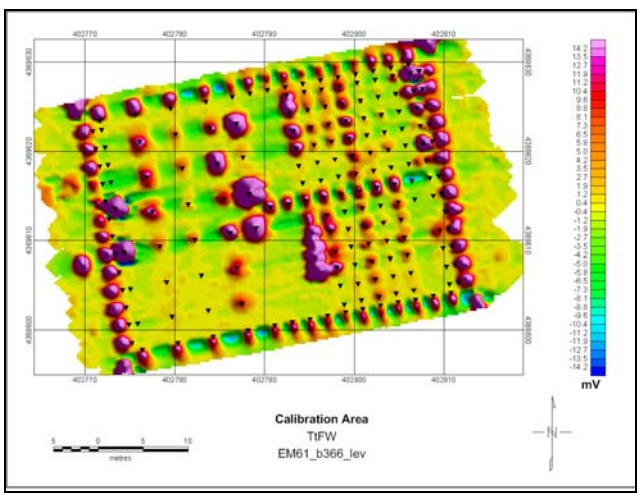
Figure 4-11. Discrimination performance for EM61 Template data, restricting the TOI subclass to UXO larger than 105mm. A) ROC curve B) Comparison of all intrinsic features versus selected C) table listing selected and intrinsic features (names ending in "_R" or "_sum" are a ratio or sum respectively)

4.3.2 Blind Test Scores

In this section, we report performance scores for blind tests that were conducted using EM61 cart data acquired by TTFW and magnetic data acquired by G-tek. These data were processed using solvers embedded into UX-Analyze and served as a shakedown for the software suite. The KNN classifier was not available when the blind test was conducted.

4.3.2.1 Labeled Data: Characterizations and Feature Selection

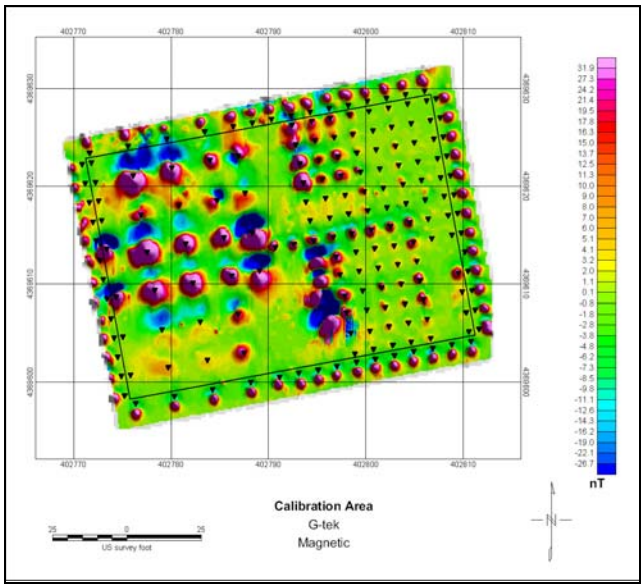
Approximately six objects of each UXO emplaced in the challenge areas were buried in the calibration area by the AEC. The objects were buried at a few depths and a few orientations, and spaced two meters apart. We inverted each target for model parameters. Some of the target signatures possessed low signal strength and could not used, while others overlapped spatially with their neighbors and could not be used. After filtering the low SNR and overlapping signatures out, we created a library of features from the remaining targets. EMI data and fit statistics area shown in Figure 4-12. Magnetic data and fit statistics are shown in Figure 4-13.



Description	Size Est. (mm)		Depth Error (mm)		Aspect Ratio		Fit Error	
20mm M55	15	2	1	0	17.3	17	30.7	0.03
40mm M385	16	3	2	0	4.5	3	26.3	0.05
M42	19	0	2	0	2.7	0	25.4	0.03
MK 118 ROCKEYE	20	1	10	0	4.4	2	33.6	0.07
16 GAGE 15cm LOOP	24	-	0	-	0.9	-	18.2	-
BDU-28	25	1	4	1	2.2	1	20.3	0.02
18 GAGE 30cm LOOP	29	-	0	-	0.7	-	13.0	-
BLU-26	30	2	3	0	3.6	2	18.1	0.03
40mm MK II	32	7	5	0	21.9	28	29.9	0.06
12 GAGE 15cm LOOP	37	7	4	5	21.1	18	27.3	0.04
57mm M86	40	3	5	0	3.5	0	14.7	0.01
8# SHOT	45	2	4	0	1.7	0	13.3	0.01
60mm M49A3	46	6	8	0	5.2	4	18.5	0.01
12# SHOT	56	-	-6	-	1.3	-	23.9	-
16 GAGE 30cm LOOP	57	-	0	-	1.0	-	16.6	-
105mm M60	67	8	15	0	1006.9	1856	26.8	0.06
81mm M374	69	1	15	0	2.5	1	19.7	0.04
105mm M456 HEAT	78	14	15	8	3.6	2	19.3	0.02
12 GAGE 30cm LOOP	80	-	0	-	2.1	-	12.1	-
30 CM STEEL PLATE	84	13	0	0	2.1	0	28.8	0.08
2.75 M230	92	15	20	0	7.0	4	14.9	0.01
T62 AT MINE	102	10	0	0	2.9	0	12.2	0.00
155mm M483A1	118	19	25	0	2.6	1	23.9	0.04
60 CM STEEL PLATE	173	11	0	0	1.9	0	17.5	0.01

Black -- mean; Blue -- standard deviation

Figure 4-12. A false-color image of the EM61 Cart data, calibration lanes, is shown on the left. Fit statistics for targets in the calibration lanes are shown on right.



Description	Size (mm)		Depth Error (mm)		Solid Angle (degrees)		Fit Error	
12 GAGE 30cm LOOP	6	-	30	-	35	-	31.6	-
16 GAGE 15cm LOOP	9	-	17	-	26	-	38.2	-
20mm M55	19	4	5	3	39	4	36.4	0.03
M42	23	1	11	6	28	11	29.8	0.01
MK 118 ROCKEYE	24	-	15	-	74	-	34.9	-
BLU-26	26	6	5	4	43	38	29.6	0.06
BDU-28	26	7	10	9	26	9	35.2	0.05
40mm MK II	45	7	9	9	24	10	32.8	0.05
60mm M49A3	48	3	7	4	50	28	35.9	0.01
8# SHOT	48	5	4	2	20	8	35.7	0.07
12# SHOT	54	-	5	-	33	-	31.7	-
57mm M86	55	6	6	4	31	7	21.1	0.00
81mm M374	59	15	7	6	40	27	29.6	0.03
105mm M456 HEAT	86	17	18	17	41	22	21.6	0.02
2.75 M230	89	2	1	0	26	14	37.1	0.03
30 CM STEEL PLATE	95	-	13	-	47	-	28.5	-
T62 AT MINE	97	1	3	2	30	4	22.0	0.03
105mm M60	119	29	12	5	46	41	21.4	0.01
60 CM STEEL PLATE	150	18	7	6	70	4	26.1	0.02
155mm M483A1	158	31	33	15	39	28	22	0.01

Black -- mean; Blue -- standard deviation

Figure 4-13. A false-color image of the magnetic data, calibration lanes, is shown on the left. Fit statistics for targets in the calibration lanes are shown on right

The approach adopted for the blind test followed a process in which all of the features from the labeled data, not just those that are intrinsic to the target, are fed into a feature selection search algorithm to determine the optimum performance. Here, we used a forward sequential search algorithm and leave-one-out methodology to determine which features produced the optimal performance for the labeled data. The features chosen are listed in Table 4-5.

Table 4-5. Selected features used by the classifiers for blind testing

	GLRT	SVM
Mag	solid angle, declination, inclination	depth, size, solid angle, declination, inclination
EMI data	depth, β2, and three orientation angles	depth, β2, and three orientation angles

4.3.2.2 Blind Test Results

We characterized and classified 1073 EMI targets and 1925 magnetic targets. The fitted features were classified using the classifiers discussed in §4.3.2.1 above. Prioritized dig lists were generated and submitted to AEC and IDA for scoring. The discrimination performance scores are shown in Table 4-6, Table 4-7, and Table 4-8.

Table 4-6. EMI Blind Scoring Results

	GLRT	SVM
TOI Retention Rate	0.82	0.57
Non-TOI Rejection Rate	0.25	0.43

Table 4-7. Magnetic Blind Scoring Results

	GLRT	SVM
TOI Retention Rate	0.87	0.60
Non-TOI Rejection Rate	0.27	0.45

Table 4-8. Depth and XY Location Error Statistics

	Depth Error (m)	Depth Error (Std Dev, m)	XY Distance Error (m)	XY Distance Error (Std Dev, m)
EMI	0.01	0.29	0.24	0.13
Mag	0.18	0.27	0.24	0.11

Discrimination statistics for discrete coherence ranges are shown in Table 4-9 thru Table 4-12. In the tables, the TOI Retention rate was calculated by dividing the number of correct ordnance declarations by the number of actual ordnance in the interval. The Non-TOI Rejection Rate for specific coherence intervals was calculated by dividing the number of correct clutter declarations by the number of actual clutter objects in the interval. Numbers are rounded to the nearest 5%. Perhaps the most significant observation that results from this analysis is that all of the methods miss-classify some ordnance regardless of coherence. Given the analysis results of the template measurements, however, this result is not unexpected.

Table 4-9. TOI Retention Rates as a function of dipole error

		Dipole Fit Error %				
TOI Retention Rate		0 to 10	11 to 14	15 to 22	23 to 32	33 to 55
EMI	GLRT	0.9	0.8	0.95	0.95	0.9
	SVM	0.65	0.5	0.55	0.6	0.4
Mag	GLRT	0	0.8	0.85	0.95	0.85
	SVM	0	0.7	0.5	0.5	0.55

Table 4-10. Non-TOI Rejection Rate as a function of dipole error

Non-TOI Rejection Rate		Dipole Fit Error % 0 to 10	11 to 14	15 to 22	23 to 32	33 to 55
EMI	GLRT	0.6	0.4	0.4	0.35	0.25
	SVM	0.55	0.6	0.5	0.6	0.45
Mag	GLRT	0.75	0.6	0.55	0.4	0.4
	SVM	0.75	0.8	0.65	0.65	0.55

Table 4-11. Depth Errors (actual minus estimated, meters)

	Dipole Fit Error % 0 to 10		11 to 14		15 to 22		23 to 32		33 to 55	
	Average	Std Dev	Average	Std Dev	Average	Std Dev	Average	Std Dev	Average	Std Dev
MAG			0.04	0.11	0.05	0.15	0.09	0.15	0.03	0.19
EMI	0.01	0.20	-0.05	0.32	-0.02	0.27	-0.07	0.51	0.21	0.61

Table 4-12. XY Distance Errors (meters)

	Dipole Fit Error % 0 to 10		11 to 14		15 to 22		23 to 32		33 to 55	
	Average	Std Dev	Average	Std Dev	Average	Std Dev	Average	Std Dev	Average	Std Dev
MAG	--	--	0.20	0.08	0.20	0.12	0.25	0.10	0.26	0.12
EMI	0.23	0.13	0.21	0.12	0.26	0.14	0.24	0.10	0.29	0.13

4.3.3 Effects of Changing the Labeled Data and Feature Selection

The blind tests used labels from the calibration area only. At the time of the blind tests, there were no other available labels. Later, additional labels for 326 objects in the Open Field and Blind Grid objects were publicly released. To investigate if the additional labels would materially impact the discrimination performance, we performed feature selection and retrained the classifiers using the new labels. Results of the new prioritizations did not improve discrimination performance (Appendix A). Given the overlapping nature of the TOI and non-TOI signatures for the EMI Cart and magnetic data, this result is not unexpected.

4.3.4 Analysis Time

The time required to quantitatively analyze individual anomalies is highly dependent on the nature of the data under consideration. The principle factors relate to the density of targets, the degree of editing required removing noise spikes or erratic positions, and the background noise due to geology, the environment, or objects on the surface. It is possible to identify, characterize, and classify hundreds of anomalies per hour if (i) the signal-to-noise ratio is high, (ii) the spatial registration of the data is consistently high, (iii) the data are well leveled, and (iv) the anomalies are separated such that they do not spatially overlap.

For this demonstration, four hours were spent preparing the Open Field magnetic data for analysis and fourteen hours for the EMI. Selecting targets and inverting for model parameters required approximately one minute per anomaly for the magnetic data and three minutes per anomaly for the EMI. Creating the anomaly plots involves two steps – creating the Oasis montaj map and exporting the map as a JPEG image. For these data, each step required approximately 0.5 seconds per anomaly (roughly 120 maps per minute).

4.3.5 Qualitative Metrics

The analysis flow consists of characterizing and classifying the anomaly. Steps taken within the classification phase include identifying anomalies, selecting the spatial footprint of the anomaly, and reviewing the results. The most tedious and labor intensive portion of the characterization phase is identifying the spatial extent of each anomaly and editing individual samples that are inconsistent. The output of the characterization phase is a detailed dig list and is then submitted to the classifier. Steps taken during the classification phase include creating or updating a library for training purposes, selecting and calling the classifier. UX-Analyze adheres to the look and feel of Oasis montaj and, as such, is easily mastered by experienced Oasis montaj users.

This demonstration was the first major shakedown of UX-Analyze. As such, we encountered bugs and logic problems that had not been exposed during development or tests with more modest data sets. During the course of this demonstration, we identified and resolved 71 bugs associated with data handling, 51 bugs associated with visualization, 15 bugs associated with the inversion routines, 22 bugs generated while adding new features and user interfaces, and 38 bugs associated with miscellaneous functionality. Due in part to the large number of data sets analyzed during this work, there are no remaining bugs that interfere with the processing flow.

End user comments were solicited from the government and regulatory community and the consulting firms providing services to the government. We also conducted a third party peer review. The goal was to solicit comments regarding the flow and functionality of UX-Analyze as well as gauge the desire for and acceptance of advanced analysis tools.

4.4 Discussion

This shakedown demonstration was instrumental in exposing a number of programming issues and bugs that were primarily related to data handling and visualization functions. These limitations were not revealed during tests conducted on small data sets.

Discriminating UXO from clutter based upon their inverted polarizabilities is inherently limited by how well the measured data can be duplicated using a dipole model source and by the extent to which the class polarization distributions overlap in feature space. The mean dipole fit error for the template EMI data was 4.7% (standard deviation of 5.1), which is quite good. By comparison, the mean fit error the production cart-mounted EMI data was 24.1% (standard deviation of 17.2. Using the fitted parameters from the template data to evaluate the separability of the TOI and non-TOI clusters, we find that the polarizabilities from non-TOI objects overlap significantly in feature space with those derived from the larger UXO. Thus, even with well positioned static data, discrimination performance at this site is limited by the distributions of the TOIs relative to the non-TOI objects. Interestingly, however, we also find that subclasses of UXO, namely 57mm and smaller UXO, can be effectively discriminated at this site. This trend is probably unique to APG and should not be extrapolated to live sites that possess extremely large numbers of small fragments and clutter items.

5 Cost Assessment

5.1 Cost Reporting

This demonstration focuses on characterizing and classifying anomalies observed in magnetic and electromagnetic data. As such, it encompasses only a small subset of costs that are typically associated with acquisition demonstrations. The relevant cost categories for data analysis demonstration are shown in Table 9.

Table 5-1. Cost categories and details

Cost Category	Sub Category	Costs
Analysis	Data Preparation	$10K*
	Anomaly Specific Characterization	$2 - $6 per anomaly
Documentation	Anomaly Summaries	$0.05 per anomaly

* this reported cost is highly variable and can deviate significantly from that reported for these data

6 Implementation Issues

6.1 End-Users Issues

The primary end-users of this data analysis technology are high-end geophysical service providers, technically orientated DoD oversight personnel, and government-sponsored researchers. In order to be successfully transitioned to the production community, this technology must not only be accepted by the Environmental Protection Agency and the Corps of Engineers, it must initiate a change in the requested deliverables regarding data analysis for UXO concerns. This program explored the methods and logic that are necessary to create a prioritized dig list.

7 References

Department of Defense, Unexploded Ordnance Response: Technology and Cost, Report to Congress, March 2001

G. Robitaille, J. Adams, C. O'Donnell, and P. Burr, 1999, *Jefferson Proving Ground Technology Demonstration Program Summary*, SFIM-AEC-ET-TR-99030, http://aec.army.mil/usaec/technology/jpgsummary.pdf.

McDonnell, P., and Karwatka, M., Standardized UXO Technology Demonstration Sites, In-progress review, 8 May 2007.

Section 349 (Public Law 105-85), Partnerships for Investment in Innovative Environmental Technologies.

Senate Report 106-50, National Defense Authorization Act for Fiscal Year 2000, May 17, 1999. Research and Development to Support UXO Clearance, Active Range UXO Clearance, and Explosive Ordnance Disposal, pages 291–293.

8 Points of Contact

ESTCP

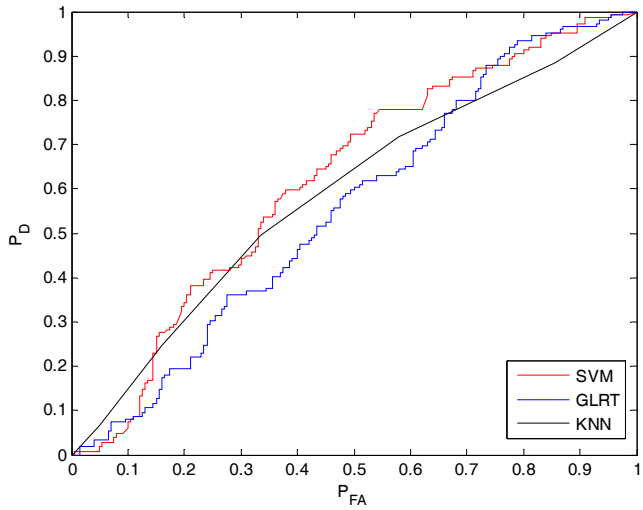
Anne Andrews	ESTCP 901 North Stuart Street Suite 303 Arlington, VA 22203	Tel: 703-696-3826 Fax: 703-696-2114 Anne.Andrews@osd.mil	Program Manager UXO Thrust Area
Jeffrey Fairbanks	HydroGeologic, Inc. 1155 Herndon Parkway Suite 900 Herndon, VA 20170	Tel: 703-736-4514 Fax: 703-471-4180 jef@hgl.com	Program Assistant UXO Thrust Area

SAIC

Dean Keiswetter	SAIC 120 Quade Drive Cary, NC 27513	Tel: 919-653-0215 Fax: 919-653-0219 keiswetterd@us.saic.com	PI
Tom Bell	SAIC 1225 Jefferson Davis Highway Suite 800 Arlington, VA 22202	Tel: 703-413-0500 Fax: 703-413-0512 bellth@us.saic.com	Co-PI
Tom Furuya	SAIC 120 Quade Drive Cary, NC 27513	Tel: 919-653-0215 Fax: 919-653-0219 furuyag@us.saic.com	Data Analyst

Appendix A: Discrimination performance for supplemental labels from the APG Open Field.

a)

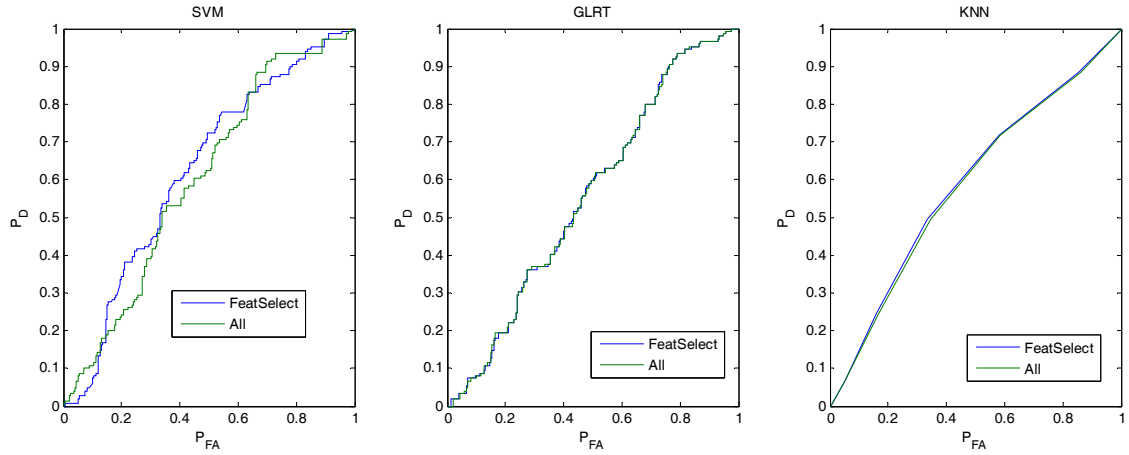


Classifier for ROC

Classifier for feature selection

	SVM	GLRT	KNN
SVM	0.621409	0.499983	0.627852
GLRT	0.597399	0.55594	0.587668
KNN	0.599564	0.554664	0.590302

Area under the ROC curve

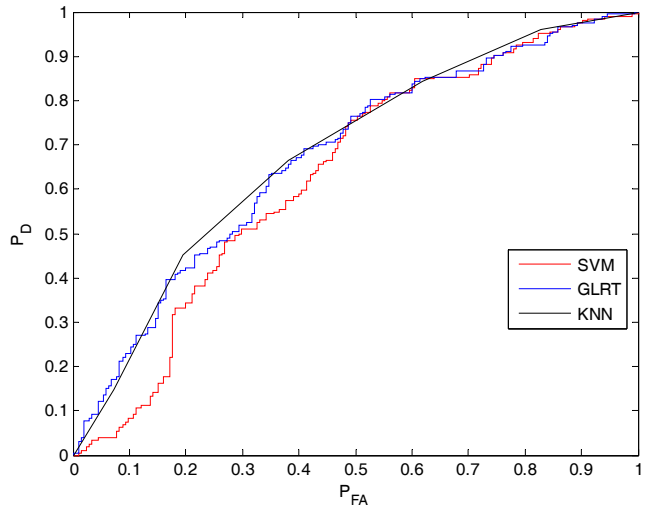


b)

c)

SVM	**GLRT**	**KNN**	**All intrinsic**
MOMENT	MOMENT	MOMENT	MOMENT
	SOLID-ANGLE	SOLID-ANGLE	SOLID-ANGLE

Figure A-1. Discrimination performance for Mag data (Open Field supplemental labels only). A) ROC curve B) Comparison of all intrinsic features versus selected C) table listing selected and intrinsic features.

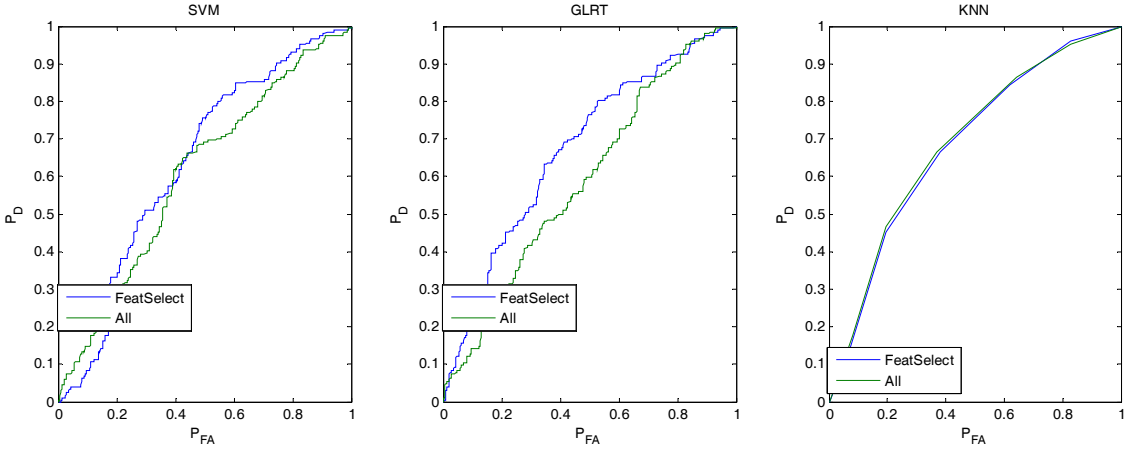


Classifier for ROC

Classifier for feature selection	SVM	GLRT	KNN
SVM	0.634122	0.609326	0.686681
GLRT	0.601172	0.673123	0.679902
KNN	0.586346	0.631564	0.682449

Area under the ROC curve

a)



b)

c)

SVM	**GLRT**	**KNN**	**All intrinsic**
B2	B1	B1	B1
B3	B3	B3	B2
Σβ	Σβ	B2/B3	B3
		Σβ	B1/B2
			B2/B3
			Σβ

Figure A-2. Discrimination performance for EMI Cart data (Open Field supplemental labels only). A) ROC curve B) Comparison of all intrinsic features versus selected C) table listing selected and intrinsic features.