

DEMONSTRATION REPORT - SWPG

Tensor Invariant Processing of Multistatic EMI Data for Target Classification

ESTCP Project MR-201310

MAY 2015

Thomas Bell
Leidos

Ryan Steigerwalt
Weston Solutions

Distribution Statement A

This document has been cleared for public release



REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE (DD-MM-YYYY) 05-31-2015		2. REPORT TYPE Demonstration Report		3. DATES COVERED (From - To) June 2014 - May 2015	
4. TITLE AND SUBTITLE Demonstration Report at Former Southwestern Proving Ground, Hope, Arkansas				5a. CONTRACT NUMBER W912HQ-13-C-0040	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) Dr. Thomas Bell				5d. PROJECT NUMBER MR-201310	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Leidos Corporation 4001 N Fairfax Drive, Suite 675 Arlington, VA 22203				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Environmental Security Technology Certification Program, 4800 Mark Center Dr Suite 17D08, Alexandria, VA 22350-3605				10. SPONSOR/MONITOR'S ACRONYM(S) ESTCP	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution unlimited					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT The objective of this project is to test and evaluate procedures for target classification in the context of the Environmental Security Technology Certification Program (ESTCP) Classification Pilot Program Live Site Demonstrations. The procedures were developed in Strategic Environmental Research and Development (SERDP) projects MR-1658 and MR-2100 [i, ii]. Classification decisions are based on two parameters easily calculated from magnetic polarizabilities of unknown targets and targets of interest. The parameters are measures of the mismatch between the strength and the shape of the respective polarizability curves.					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT UU	18. NUMBER OF PAGES 28	19a. NAME OF RESPONSIBLE PERSON Dr. Thomas Bell
a. REPORT U	b. ABSTRACT U	c. THIS PAGE U			19b. TELEPHONE NUMBER (include area code) 703-312-6288

Contents

Figures..... iii

Tables iii

Acronyms iv

1.0 Introduction..... 1

 1.1 Background..... 1

 1.2 Objective of the Demonstration 1

 1.3 Regulatory Drivers 1

2.0 Technology 2

 2.1 Technology Description..... 2

 2.1.1 TEMTADS EMI Sensors 2

 2.1.2 Processing 4

 2.2 Technology Development..... 9

 2.3 Advantages and Limitations of the Technology 9

3.0 Performance Objectives 11

 3.1 Correct Classification of TOI..... 11

 3.2 Correct Classification of Clutter 11

 3.3 Stop-Dig Threshold..... 11

 3.4 Can't Analyze 11

 3.5 Target Parameter Estimation..... 11

4.0 Site Description..... 13

5.0 Test Design 14

 5.1 Basic Experimental Design..... 14

 5.2 Validation..... 14

6.0	Data Analysis and Products	15
6.1	Preprocessing	15
6.2	Target Selection for Detection	15
6.3	Parameter Estimates	15
6.4	Classifier and Training.....	15
7.0	Performance Assessment	16
7.1	Performance Objectives	16
7.1.1	Classification: Correct Classification of TOI	16
7.1.2	Classification: Correct Classification of Clutter	16
7.1.3	Classification: Stop-Dig Threshold.....	16
7.1.4	Feature Extraction: Can't Analyze.....	16
7.1.5	Feature Extraction: Target Parameter Estimation.....	16
7.2	Overall Performance	17
8.0	Cost Assessment	18
8.1	Cost Model.....	18
8.1.1	Pre-Demonstration Development.....	18
8.1.2	Target Screening	19
8.1.3	Feature Extraction	19
8.1.4	Training.....	19
8.1.5	Classification.....	19
9.0	References.....	20
Appendix A.	Health and Safety Plan.....	A-1
Appendix B.	Points of Contact.....	B-1

Figures

Figure 1. Man-portable TEMTADS array used in the SWPG demonstration.....	3
Figure 2. TEMTADS transmitter current waveform.	3
Figure 3. Sample TEMTADS cued data set. Ordinate is signal in mV normalized by Tx current, abscissa is time gate.....	4
Figure 4. Singular value decomposition of man-portable TEMTADS array data. Left (anomaly 1032) is a partial rifle grenade, Right (anomaly 1030) is a small piece of wire.....	5
Figure 5. Principal axis polarizabilities for a 57 mm projectile (left) and a horseshoe (right).	6
Figure 6. Classification parameter distributions for Camp Beale man-portable TEMTADS array demonstration. Top: scatter plot of size and shape mismatch parameters. Middle: cumulative distribution of net TOI mismatch. Bottom: decision metric values rank ordered from most like TOI to least like TOI. Values for TOI items plotted in red, clutter items in blue.	8
Figure 7. Decision metric distributions for a data set from the Pole Mountain demonstration (see text).....	10
Figure 8. ROC curve for SWPG classification demonstration.	17

Tables

Table 1. Performance objectives.....	12
Table 2. Demonstration costs.....	18

Acronyms

EMI	Electromagnetic Induction
ESTCP	Environmental Security Technology Certification Program
FOM	Figure of Merit
GPS	Global Positioning System
GSV	Geophysical System Verification
IDA	Institute for Defense Analyses
ISO	Industry Standard Object (seeded pipe section)
IVS	Instrument Verification Strip
MTADS	Multisensor Towed Array Detection System
MR	Munitions Response
RF	Recovery Field
ROC	Receiver Operating Characteristic
Rx	Receive
SERDP	Strategic Environmental Research and Development Program
SNR	Signal to Noise Ratio
SWPG	Southwestern Proving Ground
TEM	Transient Electromagnetic
TEMTADS	Transient Electromagnetic MTADS
Tx	Transmit

1.0 INTRODUCTION

The objective of this project is to test and evaluate procedures for target classification in the context of the Environmental Security Technology Certification Program (ESTCP) Classification Pilot Program Live Site Demonstrations. The procedures were developed in Strategic Environmental Research and Development (SERDP) projects MR-1658 and MR-2100 [1, 2]. Classification decisions are based on two parameters easily calculated from magnetic polarizabilities of unknown targets and targets of interest. The parameters are measures of the mismatch between the strength and the shape of the respective polarizability curves.

1.1 BACKGROUND

The characterization and remediation activities conducted at Department of Defense sites contaminated with unexploded ordnance (UXO) using 1990's and early 21st century technology often yield unsatisfactory results and are too expensive. In part, this is due to the inability of that technology to distinguish between UXO and non-hazardous items. Field experience has shown that when using the old technology over 90% of objects excavated during the course of remediation can be non-hazardous clutter.

SERDP and ESTCP have developed and tested several purpose-built multi-axis electromagnetic induction (EMI) sensor array systems for classifying buried objects at munitions response sites. They have also invested in developing new processing procedures optimized for this new generation of EMI sensors. This demonstration serves to evaluate the performance of procedures developed in SERDP projects MR-1658 and MR-2100.

1.2 OBJECTIVE OF THE DEMONSTRATION

The objective of this demonstration is to demonstrate the classification performance of the procedures developed in SERDP projects MR-1658 and MR-2100 using data collected by NAEVA Geophysics with the man-portable transient EMI (TEM) array (nicknamed "TEMTADS" [3]) at the former Southwestern Proving Ground (SWPG) near Hope, Arkansas. The data collection followed the approach outlined in demonstration plans from the ESTCP Program Office [4] and Weston Solutions, Inc. [5].

1.3 REGULATORY DRIVERS

The ESTCP has assembled an Advisory Group to address the regulatory, programmatic and stakeholder acceptance issues associated with the implementation of classification in the Munitions Response (MR) process. Details can be found in their guide to implementing advanced classification on munitions response sites [6].

2.0 TECHNOLOGY

The data collection technology and approach are discussed in the Weston Solutions, Inc. Demonstration Plan [5]. NAEVA used the man-portable TEMTADS array to collect survey mode data which was then used to identify metallic anomalies within the study area. The survey data were analyzed to produce a list of anomalies considered to have the potential to be targets of interest (TOI). TOI include intact munitions items and pipe sections which simulate the EMI signatures of munitions items and which are implanted at the site for quality control and assurance purposes. They then parked the array over each of these anomalies in turn and collected “cued” data to be used for target classification. In this demonstration we processed the cued data to determine the likelihood that the anomaly is actually due to a TOI.

The processing and analysis procedures were developed in SERDP projects MR-1658 and MR-2100 for use with the advanced EMI arrays developed by SERDP and ESTCP for target classification. Classification typically involves comparing principal axis polarizabilities calculated from EMI data collected over an unknown target with those of known TOI [6]. The classification algorithm used here exploits the fact that an object’s polarizability is a product of two factors: the volume of the object and a tensor whose eigenvalues depend only on the shape and composition of the object. Confronted with an unknown target, we compare its apparent size and EMI “shape” with the sizes and shapes of TOI. Classification is based on thresholding a figure of merit (FOM) parameter that is a weighted sum of parameters quantifying the mismatches in the EMI size and shape of the target relative to the TOI. For multiple TOI, the FOM is minimized over the set of TOI. This basic algorithm is also used in cluster analysis to identify unexpected munitions. In this case each target is compared against all others to find groups which have similar EMI size and shape.

2.1 TECHNOLOGY DESCRIPTION

2.1.1 TEMTADS EMI Sensors

The man-portable TEMTADS array used in the SWPG demonstration consists of four single-axis transmit (Tx) coils and four three-axis receive (Rx) cubes arranged in a 2x2 array of Tx/Rx pairs. The picture on the left in Figure 1 shows one of the large (35 cm square by 8 cm high) Tx coils and one of the 8 cm Rx cubes which fits inside the foam core of the Tx coil. The middle picture shows three of the Tx/Rx pairs set into the plastic array enclosure. The centers of the Tx/Rx pairs are spaced 40 cm apart. The picture on the right shows the assembled array with its Global Positioning System (GPS) antenna. The array is mounted on a cart with the coils 20 cm above the ground.



Figure 1. Man-portable TEMTADS array used in the SWPG demonstration.

Currents through the Tx coils illuminate a target in the ground under the array with an alternating bipolar magnetic field (primary field), which excites eddy currents in the target. Figure 2 shows the Tx current waveform for a block time of 0.9 s with nine repeats of the basic bipolar cycle within the block. This was the pattern used for the cued data at SWPG. The Rx coils measure the decay of the secondary magnetic field from the eddy currents during the intervals between the alternating pulses of positive and negative Tx current. The measured responses are averaged over the block after inverting those from the negative current pulses. A sequence of 18 of these blocks was collected over each target and the net responses from the 18 blocks were stacked (averaged) to produce the recorded EMI response for each of the 48 possible Tx/Rx combinations (four transmitters and each of the three axes of the four receivers).

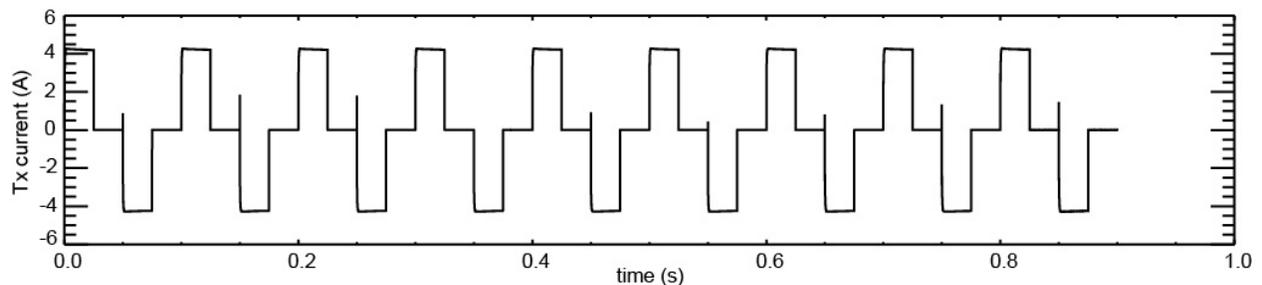


Figure 2. TEMTADS transmitter current waveform.

Secondary field data are recorded for 121 time gates spaced logarithmically out to 25 ms after the primary field cutoff (25 ms is just the time interval between the end of one current pulse and the beginning of the next). Figure 3 is an example of the TEMTADS data collected over a target. Background response has been removed from the signals as described in section 2.1.2 below. Each panel corresponds to a different Tx/Rx pair, and different colors are used to show the responses for the different Rx cube axes. The ordinate (vertical axis) scale is the background-subtracted signal in mV normalized by the peak Tx current and the abscissa (horizontal axis) scale is time gate.

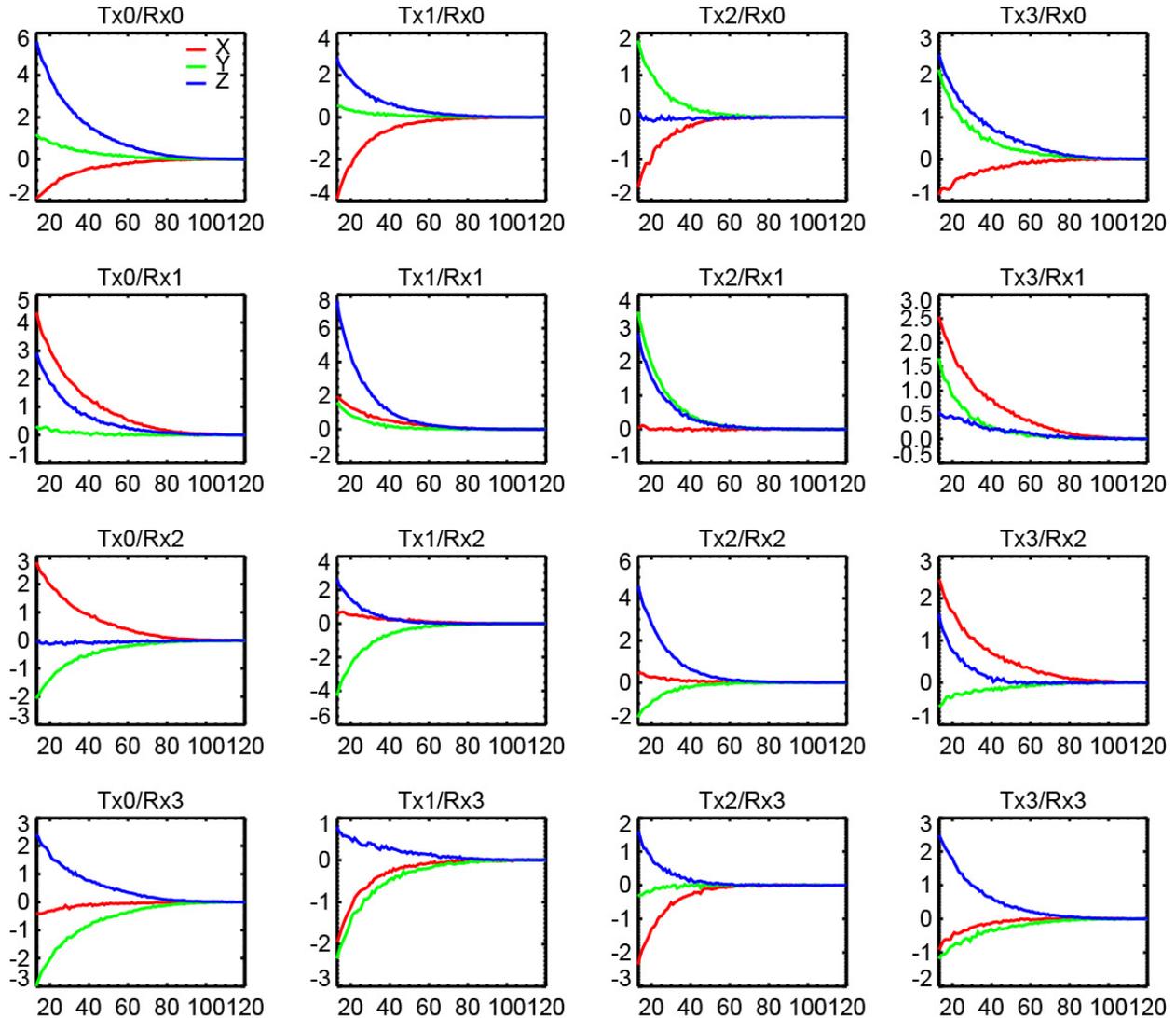


Figure 3. Sample TEMTADS cued data set. Ordinate is signal in mV normalized by Tx current, abscissa is time gate.

2.1.2 Processing

As recorded the data include a substantial background response caused primarily by electronic ring-down following the primary field shutoff at the end of each current pulse. This is removed from the data by subtracting background shots taken over nearby, nominally target-free ground. We skip the first 12 gates because the very early time ring-down effects overwhelm the target response. The plots in Figure 3 show background-subtracted signals.

The first stage in the processing of background-subtracted data is singular value decomposition. This is used primarily as a screening tool to determine whether or not the signal is strong enough that we can calculate the target's principal axis polarizabilities, which will subsequently be used

to classify the target as TOI or clutter. With a 4x12 matrix of Tx/Rx combinations we can resolve four singular value components. Three degrees of freedom (principal axis polarizabilities) are needed to represent a target. With a good strong signal from a single target three of the singular value components will correspond to linear combinations of the polarizabilities while the fourth will correspond to the measurement noise. For targets whose signals are too weak to determine the polarizabilities, all will drop down to the noise level. Figure 4 shows examples from data collected in June 2013 at Marine Corps Base Quantico. In the plot on the left there are three singular values above the noise level out to a few ms. These data can be inverted and classified. The recovered target for this anomaly (1032) was a partial rifle grenade. In the plot on the right the singular values are buried in the noise. These data cannot be inverted. The recovered target for this anomaly (1030) was a small piece of wire.

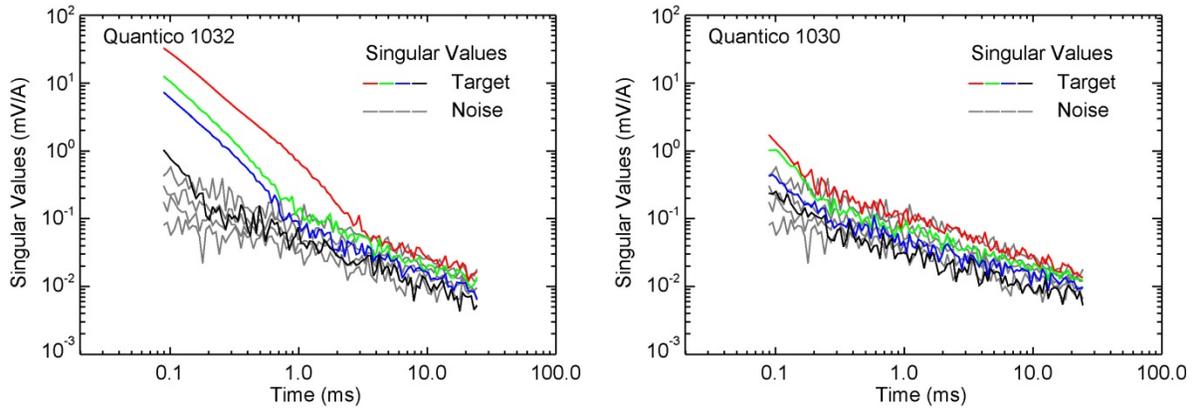


Figure 4. Singular value decomposition of man-portable TEMENTADS array data. Left (anomaly 1032) is a partial rifle grenade, Right (anomaly 1030) is a small piece of wire.

If the data can support inversion, principal axis polarizabilities are calculated using a signal to noise ratio (SNR) weighted [7] inversion algorithm. The principal axis polarizabilities are the basis for classification. Figure 5 Shows principal axis polarizabilities for two quite different objects: a 57 mm projectile (left) and a horseshoe (right) encountered at the Remington Woods site in Bridgeport, CT. The objects are similar in size but have quite different shapes. Taken together the sets of three principal axis polarizabilities are quite different for the two objects.

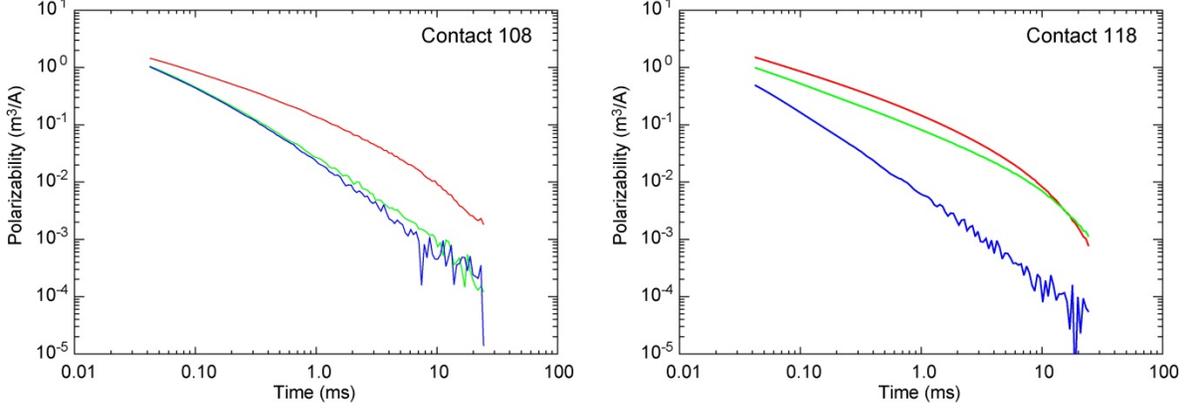


Figure 5. Principal axis polarizabilities for a 57 mm projectile (left) and a horseshoe (right).

Classification exploits these differences. Classification is a matter of deciding whether the object’s polarizabilities are munitions-like or clutter-like. Library matching methods employing various procedures to compare polarizabilities of unknown targets with those of TOI items are commonly used for classification. Ours exploits the fact that an object’s polarizability tensor $\beta_{ij}(t) = V\alpha_{ij}(t)$ is a product of two factors: the volume V of the object and a tensor $\alpha_{ij}(t)$ whose eigenvalues $\alpha_i(t)$, $i = 1, 2, 3$ are determined by the shape and composition of the object. Confronted with an unknown target, we compare its apparent size and EMI “shape” with the sizes and shapes of the TOI.

Given the set (spanning three axes and N time gates) of principal axis polarizabilities β_0 for a TOI and the set of principal axis polarizabilities β for an unknown target, we calculate a size ratio s as

$$s = \text{median} \left(\frac{\sqrt[3]{\beta}}{\sqrt[3]{\beta_0}} \right)$$

where the median is taken over all axes and time gates for the polarizabilities are above some threshold level which reflects the expected inversion noise. If a significant fraction (typically 25-50%) of the available polarizability terms are below this threshold, then the target is put in the “can’t analyze” category. We define the size ratio in terms of the cube root of polarizability because polarizability scales with target volume (linear dimensions cubed).

The size mismatch parameter Δ_{size} is defined as

$$\Delta_{\text{size}} = \log(s)$$

which is equal to zero if the EMI sizes of the target and the reference TOI are the same, The shape mismatch parameter Δ_{shape} is determined by comparing the unknown target’s polarizability with the reference polarizability scaled by the size ratio

$$\Delta_{shape} = \frac{\sum |\sqrt[3]{\beta} - s\sqrt[3]{\beta_0}|}{\sum \sqrt[3]{\beta}}$$

in which the sums are over all terms with β above the noise level. Optionally the three principal axis polarizabilities can be assigned different weights W_i in calculating the shape mismatch. For each target, size and shape mismatch parameters are calculated for each TOI. By combining the size and shape mismatch parameters we can define a net TOI mismatch parameter as

$$TOI \text{ Mismatch} = \min_{TOI} \{ |\Delta_{size}| + k \log(\Delta_{shape}) \}.$$

We have found that using a parameter value $k \approx 0.3$ gives the best classification performance. Low values of the TOI mismatch indicate a good match to both the size and the shape of the TOI. Minimizing the parameter over the set of TOI finds the best match to any TOI.

The TOI mismatch parameter typically runs between about -1 and 1, with TOI having the lowest values (best match of target polarizability strength and decay curve shapes to library polarizabilities) and clutter having the highest values (poor match to TOI polarizabilities). Figure 6 shows the distributions of the size and shape parameters (top plot) and the cumulative distribution of the net TOI mismatch (middle plot) for the man-portable TEMTADS array at the Camp Beale classification demonstration. Values for targets identified as TOI using the post-test ground truth are plotted in red and those for clutter items in blue.

Classification is based on thresholding a decision metric related to the TOI mismatch. For the sake of consistency with conventions used by other demonstrators (i.e., that TOI have large values of the decision metric and clutter items have small values) we define the decision metric as one over the antilog of the TOI mismatch, which works out to be

$$Decision \text{ Metric} = \max_{TOI} \left\{ \min \left(s, \frac{1}{s} \right) (\Delta_{shape})^{-k} \right\}.$$

The first term in the curly brackets (s or s^{-1} , whichever is smaller) equals one when the EMI size of the target matches the TOI. Otherwise the value of the decision metric is reduced by the extent that the target size differs from the TOI size. The second term is larger when the polarizability shapes match well and smaller when they do not. The bottom plot in Figure 6 shows the decision metric values rank ordered from most like TOI to least like TOI. Again, TOI values are shown in red and clutter values in blue. There is a distinct bend or slope break in the distribution as we go from TOI to clutter, followed by a gradual decline as we chew through the clutter items. We see similar patterns in the decision metric distributions obtained by re-processing data from other of the ESTCP Classification Pilot Program Live Site Demonstrations, leading us to conclude that with good quality control the stop-dig threshold may be set at the end of the slope break. As a practical matter the threshold has to be set low enough to capture those TOI which for some reason do not match the library specimens as well as most, and so setting the stop-dig point tends to be a bit of an art.

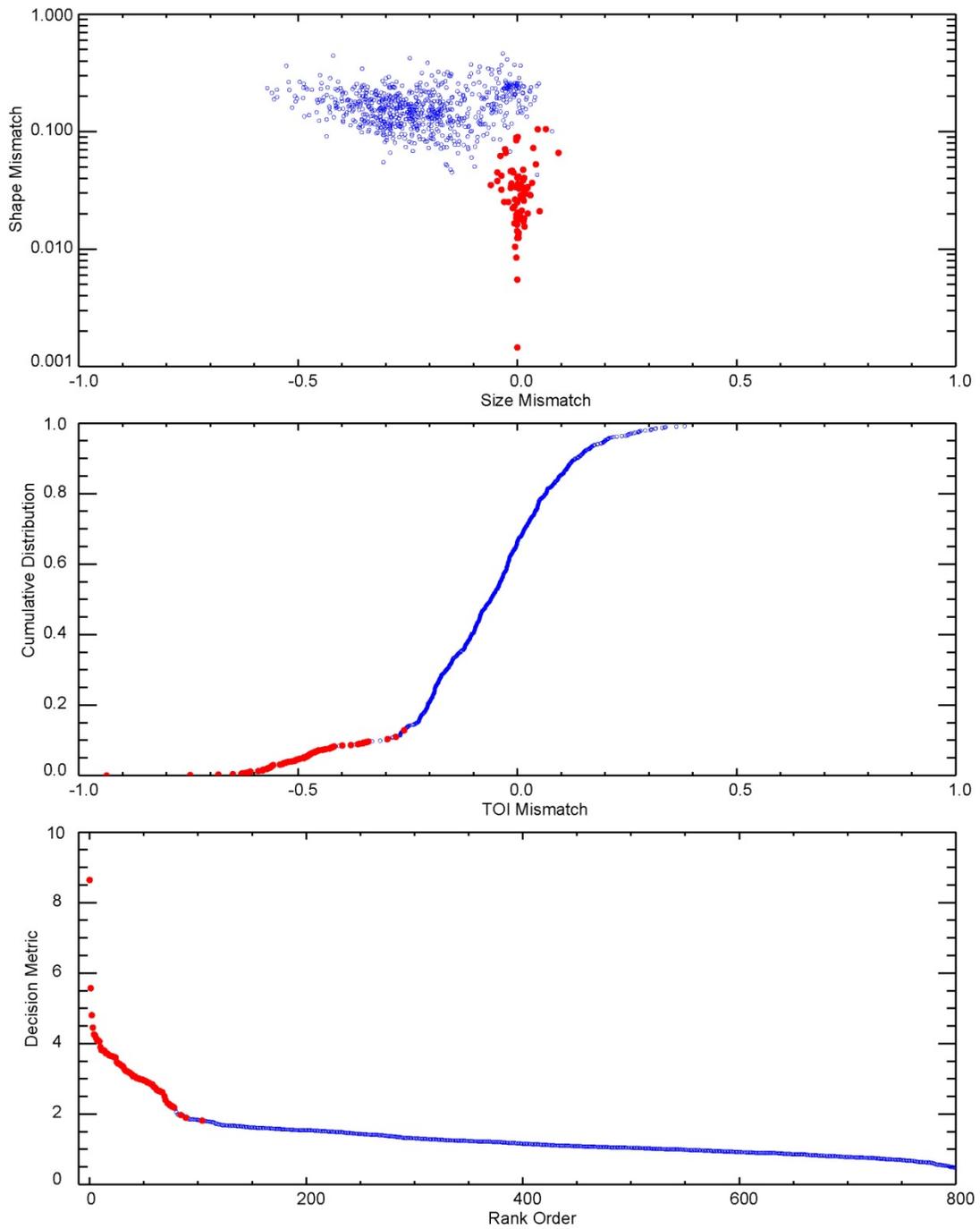


Figure 6. Classification parameter distributions for Camp Beale man-portable TEMTADS array demonstration. Top: scatter plot of size and shape mismatch parameters. Middle: cumulative distribution of net TOI mismatch. Bottom: decision metric values rank ordered from most like TOI to least like TOI. Values for TOI items plotted in red, clutter items in blue.

2.2 TECHNOLOGY DEVELOPMENT

The development work for this project was done under SERDP projects MR-1658 and MR-2100 and is documented in references [1, 2].

2.3 ADVANTAGES AND LIMITATIONS OF THE TECHNOLOGY

The advantages of this technology are both quantitative and qualitative. Re-processing data from the recent Camp Beale demonstration using this approach produced a Receiver Operating Characteristic (ROC) curve [6] which rises more rapidly and hits the 100% TOI recovered level with 50% fewer clutter digs beyond the training set than the ROC from conventional processing. Improved classification performance improves munitions response efficiency. The procedure operates in an intuitive and easily visualized feature space. It is transparent, objective and easily automated. All of this is likely to facilitate transition to production work and ease regulatory acceptance.

The decision metrics in common use at this time were described at the 2015 ESTCP Demonstrators Meeting [8]. The Black Tusk Geophysics decision metric is similar to our shape mismatch parameter:

$$Decision\ Metric\ (BTG) = \sum_{axes} W_i \frac{\sqrt{\sum_{gates} (\beta_{i,0}^\gamma - \beta_i^\gamma)^2}}{\text{mean}(\beta_i)}$$

with $\gamma = 0.1$. The only significant differences between our decision metric and the Black Tusk decision metric are that we explicitly separate out a scale factor and they flatten the polarizabilities more than we do ($\gamma = 0.1$ vs. $1/3$). As noted previously, we use $\gamma = 1/3$ to transform the volume scaling of the polarizability into a size or linear dimension scaling. We explicitly pull out the scale factor to accommodate the fact that small errors in the target depth estimate obtained during data inversion can significantly alter the magnitude of the calculated polarizability without significantly altering the shape of the polarizability curves.

The decision metric employed in UX-Analyze, originally constructed for the early TEMTADS demonstrations, is more complicated than the others. It compares the polarizabilities of an unknown target with each library entry based on 3 criteria: the amplitude of the primary polarizability β_1 ($= L_1$ in the formula below) and two shape parameters based on the secondary polarizabilities, $L_2 = \beta_2/\beta_1$ and $L_3 = \beta_3/\beta_1$:

$$Decision\ Metric\ (UXA) = \sum_{axes} W_i \exp\left(-\frac{1}{2\sigma} \left[\sum_{gates} |L_i - L_{i,0}| \right]^2\right).$$

The difference in the values is computed at all time gates save those where the values are negative. The differences are plugged into a Gaussian with standard deviation σ derived by

examining the variability in the amplitude and shape parameters for a large number of objects with known ground truth [9]. The decision metric is a weighted average of the results from the three different criteria and ranges from 0 (worst possible fit) to 1 (perfect fit). Although in form it resembles a probability measure it is not intended to be such. The semblance arises only because the Gaussian function provides a simple way to map the parameters onto the range $[0,1]$.

Figure 7 (after [8]) shows distributions of the three decision metrics for a common data set from the Pole Mountain demonstration. Note the resemblance of the first two to the lower plot in Figure 6. These distributions are referred to as L-curves, and the stop-dig threshold is usually set somewhere beyond the point where the L-curve has completed its initial drop off and settled into its final slow decay (somewhere beyond 200 digs in this case). This break is not so apparent with the UXA decision metric.

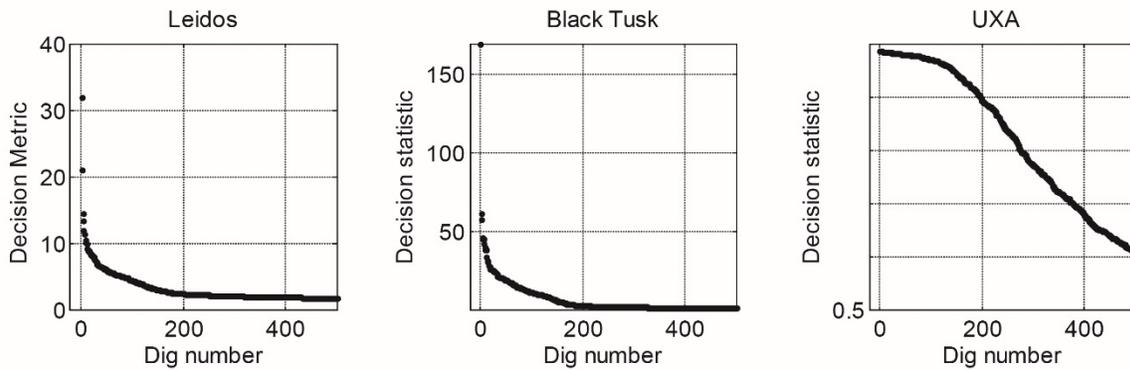


Figure 7. Decision metric distributions for a data set from the Pole Mountain demonstration (see text).

3.0 PERFORMANCE OBJECTIVES

The performance objectives for this demonstration are summarized in Table 1. The goal is to correctly classify all TOI as TOI and as many as possible clutter items as clutter. The first three objectives refer to the classification part of the demonstration with the first two referring to evaluation of the results in a retrospective analysis using scoring results from the Institute for Defense Analysis (IDA) and the third addressing how well we are able to specify the correct stop-dig (all TOI recovered) threshold in advance. The final two objectives refer to target feature extraction. All of the performance objectives were met.

3.1 CORRECT CLASSIFICATION OF TOI

There were six TOI in the cued demonstration area (four 2"x8" seeded pipe sections, one 37mm projectile and one 75mm projectile). All of the TOI were correctly classified as such in the final dig list submitted to ESTCP for scoring. The objective was to correctly identify all TOI.

3.2 CORRECT CLASSIFICATION OF CLUTTER

There were 492 clutter items. At the point where all TOI had been identified only 17 clutter items had been marked for digging, leaving 475 clutter items (96.5%). The objective was to reduce clutter digs by >85% while retaining all TOI

3.3 STOP-DIG THRESHOLD

At the stop-dig threshold all TOI were correctly identified and 18 clutter items had been marked for digging, leaving 474 clutter items (96.3%). The objective was to set the stop-dig threshold so that all TOI were identified and clutter digs had been reduced by >85%.

3.4 CAN'T ANALYZE

The objective was that reliable target parameters would be estimated for >95% of the anomalies, leaving <5% in the "Can't Analyze" category. No targets were designated "Can't Analyze" in the final dig list

3.5 TARGET PARAMETER ESTIMATION

This objective specified that the accuracy of target parameters extracted from data collected over the seed items result in <10% variation in the size estimates ($\pm 5\%$) and <10% shape mismatch. There were four 2"x8" standard pipe sections commonly referred to as industry standard objects or ISOs [10] seeded in the cued demonstration area. The overall spread in size estimates was 5.5% and the maximum shape mismatch was 3.2%.

Table 1. Performance objectives

Objective	Metric	Data Required	Success Criteria	Results
Maximize correct classification of TOI	Number of TOI retained	<ul style="list-style-type: none"> • Ranked anomaly list • Scoring report from IDA 	Correctly classify all TOI	No mis-classified TOI
Maximize correct classification of clutter	Number of clutter digs eliminated	<ul style="list-style-type: none"> • Ranked anomaly list • Scoring report from IDA 	Reduction of clutter digs by >85% while retaining all TOI	Reduction of clutter digs by 96.5% while retaining all TOI
Specification of Stop-dig threshold	Probability of correct TOI classification and number of clutter digs at threshold	<ul style="list-style-type: none"> • Stop-dig threshold • Scoring report from IDA 	Stop-dig threshold to achieve classification objectives	100% of TOI and 96.3% of clutter correctly identified at threshold
Minimize number of anomalies that cannot be analyzed	Number of anomalies classified as “Can’t Analyze”	Target parameters extracted from data collected over target	Reliable target parameters estimated for >95% of anomalies	No “Can’t Analyze” targets
Correct estimation of target parameters	Accuracy of estimated target parameters for seed items	Target parameters extracted from data collected over seeds	<ul style="list-style-type: none"> • Size ratio variation <0.1 • Shape mismatch <0.1 	<ul style="list-style-type: none"> • Overall size mismatch spread 0.055 • Maximum shape mismatch 0.032

4.0 SITE DESCRIPTION

The demonstration site is located within Recovery Field (RF) 15 of the former Southwestern Proving Ground in the southwestern corner of Arkansas. RF 15 is in open farmland and is relatively even grade across the site. Known and suspected munitions types in RF 15 include 20mm, 37mm, 40mm, 57mm, 75mm, 76mm, 90mm, 105mm and 155mm projectiles and 81mm mortars. Site maps and additional relevant information are included in the Weston Solutions, Inc. Demonstration Plan [5].

5.0 TEST DESIGN

5.1 BASIC EXPERIMENTAL DESIGN

The basic experimental design is described in the Weston Solutions, Inc. Demonstration Plan [5]. Our demonstration entails analysis of data collected over nominally 500 anomalies in a roughly 0.16 Ha (0.4 acre) area of RF-15. These data were collected by NAEVA Geophysics personnel.

The ESTCP Program Office coordinated data collection activities and provided us with survey and/or cued data over selected anomalies. We processed the data to extract target parameters which were then be passed to our classification routines. After training on data from previous demonstrations, data from the instrument verification strip (IVS) and test pit, and a limited amount of site-specific ground truth, the classification routines were used to produce ranked anomaly lists. We requested ground truth data on 20 targets for training to identify possible unforeseen munitions types. None were found. At the conclusion of this training, we submitted a ranked anomaly list. According to ESTCP instructions, the list is structured such that all anomalies for which training labels were requested are placed at the top of the list. Then, the anomalies for which we were not able to extract meaningful parameters would be listed (we had none). Following these “can’t extract reliable parameters” anomalies, the list is ordered from the item we are most confident is TOI through the item we are most confident is not TOI. The dig list was scored by IDA with emphasis on the number of items that are correctly labeled non-hazardous while correctly labeling all TOI. The primary objective of the demonstration is to assess how well we are able to order our ranked anomaly list and specify the threshold separating high confidence clutter from all other items.

5.2 VALIDATION

At the conclusion of data collection activities, all anomalies on the master anomaly list assembled by the Program Office were excavated. Each item encountered was identified, photographed, its depth measured, its location determined using cm-level GPS, and the item removed if possible. This ground truth information was used to validate the objectives listed in Section 3.0

6.0 DATA ANALYSIS AND PRODUCTS

6.1 PREPROCESSING

The data were pre-processed by the data collection demonstrators (NAEVA Geophysics) and provided to us by the ESTCP Program Office. We applied appropriate background subtraction prior to feature extraction (parameter estimation).

6.2 TARGET SELECTION FOR DETECTION

Targets were selected by the data collection demonstrators (NAEVA Geophysics) and a master target list and corresponding cued data were distributed by the ESTCP Program Office.

6.3 PARAMETER ESTIMATES

Target parameters (polarizabilities) were extracted using SNR weighted dipole inversion augmented by other relevant processing techniques developed in SERDP projects MR-1658 and MR-2100.

6.4 CLASSIFIER AND TRAINING

Classification is based on thresholding our decision metric

$$Decision\ Metric = \max_{TOI} \left\{ \min \left(s, \frac{1}{s} \right) (\Delta_{shape})^{-k} \right\}$$

where s is our unknown-to-library target size ratio and Δ_{shape} is our shape mismatch parameter, which are determined by comparing an unknown target's polarizabilities with TOI polarizabilities. We set the parameter $k = 0.3$. The stop-dig threshold was set based on experience with previous demonstration data sets (Pole Mountain, Camp Beale, Camp Spencer), observed parameter spreads with training data and polarizability errors bars from the inversion processing.

Training data were chosen based on a cluster analysis in which each target's polarizabilities are compared with those of all other targets using the size and shape matching parameters. This procedure identifies groups of targets of similar size and shape, including possible unexpected munitions items which would then become TOI. Training data are also used in selecting the stop-dig threshold. Following training we produced a ranked anomaly list with a threshold that corresponds to those items that should be investigated in the first round of intrusive work. Normally the ranked anomaly list would be refined based on feedback regarding seeds and first-round intrusive work and process iterated until we are satisfied with our classification results, at which point a final ranked dig list is submitted for scoring. In this case we were satisfied with the results of the first round and used the initial ranked anomaly list for our final submission.

7.0 PERFORMANCE ASSESSMENT

7.1 PERFORMANCE OBJECTIVES

There were five performance objectives (see Table 1). The first three refer primarily to the performance of the classification algorithm, while the others refer to the performance of the target feature extraction algorithm. All of the performance objectives were met in this demonstration.

7.1.1 Classification: Correct Classification of TOI

The objective was to correctly identify all TOI. Performance evaluation is determined directly from the IDA scoring of our final ranked anomaly (dig) list. There were six TOI in the cued demonstration area (four 2"x8" seeded pipe sections, one 37mm projectile and one 75mm projectile). All of the TOI were correctly classified as such in the final dig list submitted to ESTCP for scoring by IDA.

7.1.2 Classification: Correct Classification of Clutter

The objective was to reduce clutter digs by >85% while retaining all TOI. Performance evaluation is determined directly from the IDA scoring of our final ranked anomaly (dig) list. There were 492 clutter items. At the point where all TOI had been identified only 17 clutter items had been marked for digging, leaving 475 clutter items (96.5%).

7.1.3 Classification: Stop-Dig Threshold

At the stop-dig threshold all TOI were correctly identified and 18 clutter items had been marked for digging, leaving 474 clutter items (96.3%). The objective was to set the stop-dig threshold so that all TOI were identified and clutter digs had been reduced by >85%.

7.1.4 Feature Extraction: Can't Analyze

This is as much a function of data quality and signal strength relative to noise as it is of the quality of the algorithm used for inverting the data to extract target features (principal axis polarizabilities). The objective was that reliable target parameters would be estimated for >95% of the anomalies, leaving <5% in the "Can't Analyze" category. We did not distinguish between "can't analyze" and "can't extract reliable parameters." Overall the data quality was very good and it turned out that no targets had to be designated "Can't Analyze" in the final dig list.

7.1.5 Feature Extraction: Target Parameter Estimation

By comparing the target features (polarizabilities) of the seed items one against the other we can get a quantitative measure of the accuracy of the feature extraction process for these data. We use the size and shape mismatch parameters described in section 2.1.2 to compare the polarizabilities of the seed items with each other. The objective specified that the accuracy of

target parameters extracted from data collected over the seed items result in <10% variation in the size estimates and <10% shape mismatch relative to the nominal polarizability for the seeded object. There were four 2”x8” medium sized ISOs (standard pipe sections) seeded in the cued demonstration area. The overall spread in size estimates was 5.5% and the maximum shape mismatch was 3.2% relative to the nominal polarizability for a medium sized ISO.

7.2 OVERALL PERFORMANCE

Overall performance in the ESTCP classification demonstrations is summarized by the ROC curve. This is a plot of the number of TOI items recovered as a function of the number of clutter digs as we move through the dig list [6]. The ROC curve for this demonstration produced as part of the IDA scoring report is shown in Figure 8. The dashed portion corresponds to the anomalies selected for training data and the blue dot is our stop-dig point (see section 7.1.3 above).

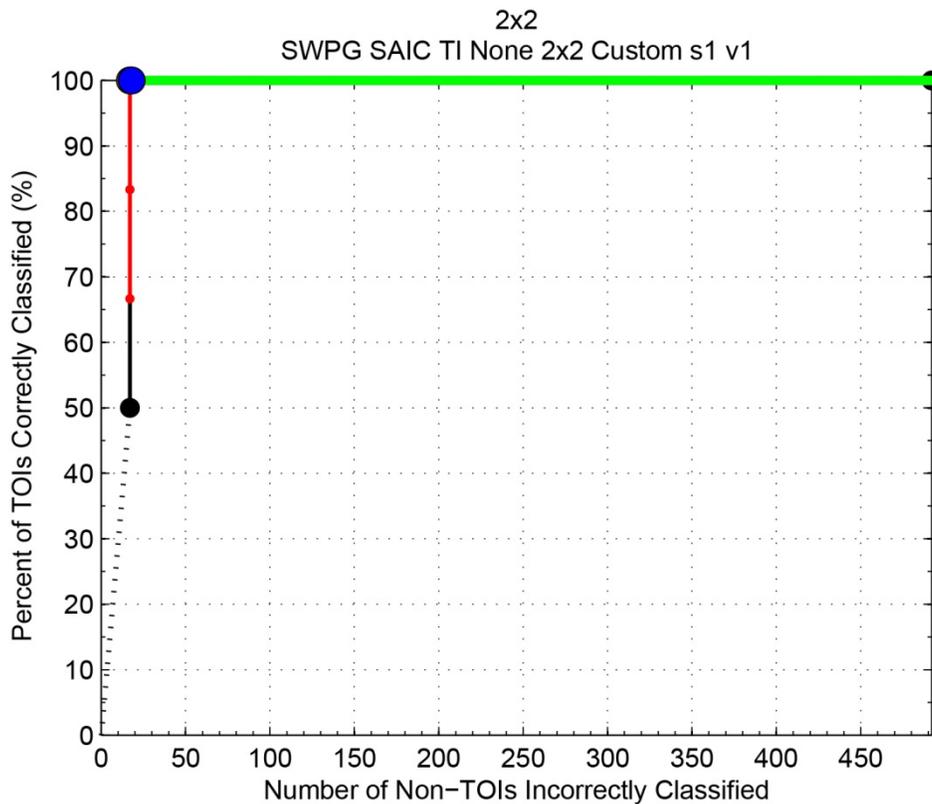


Figure 8. ROC curve for SWPG classification demonstration.

8.0 COST ASSESSMENT

8.1 COST MODEL

The cost elements for this demonstration are listed in Table 2. They are discussed below in the corresponding sub-sections. Computations were done on a 2011 Dell Latitude E6420 business laptop computer with a 2.8GHz Intel Core i7-2640M processor.

Table 2. Demonstration costs.

Cost Element	Data Tracked During Demonstration	Estimated Costs
Pre-demonstration Development	Not tracked	Not tracked
Target Screening	Time required to perform singular value decomposition on the data and determine suitability for calculating principal axis polarizabilities.	2 hr
Feature Extraction	Time required to calculate principal axis polarizabilities.	5.15 hr
Training	Time required to establish appropriate TOI libraries	32 hr
Classification	Time required to apply classification algorithms and produce dig list.	0.5 hr

8.1.1 Pre-Demonstration Development

As directed by the ESTCP Program Office, pre-demonstration development focused on analysis of multiple targets because of the likelihood that demonstrations would involve a significant number of anomalies corresponding to multiple sources. Several weeks of analysis effort were spent comparing the performance of processing techniques for multiple-target anomalies. The two basic approaches are the UX-Analyze multisolver [11] and standard dipole inversion using more than one dipole source term. On comparing the results of the two approaches on several data sets we concluded that the latter gave more consistent and reliable results and it was adopted as our basic technique for analyzing multiple target anomalies in the demonstration.

8.1.2 Target Screening

Singular value decomposition was used for target screening. This is a very quick computation: roughly ten seconds to perform singular value decomposition and produce postscript plots for the 498 anomalies. The plots are visually screened to identify any spurious anomalies and then held for reference during the classification stage. As described above in Section 2.1.2 the plots can indicate that the target response is in the noise or that the response includes contributions from multiple targets. The estimated cost for this element is two hours, which includes 20 minutes for the original visual inspection of the plots to identify unusual anomalies (there were none) and another 100 minutes for referring back to the plots for those anomalies that required extra attention during the classification stage.

8.1.3 Feature Extraction

The total time for feature extraction was 5.15 hours. This included 9 minutes for single dipole inversions, 63 minutes for two dipole inversions and 237 minutes for three dipole inversions. The man-portable TEMTADS does not support reliable higher-order inversions with cued data than the three dipole fit.

8.1.4 Training

In the training phase we assemble a suitable TOI library to be used for classification. Several potential TOI items (20mm, 37mm, 40mm, 57mm and ISO) were measured by NAEVA Geophysics in the test pit and IVS strip at SWPG. Polarizabilities for these items were included in our classification library. We have kept a catalog of polarizabilities of munitions items and ISO measured throughout the course of the various TEMTADS projects, and as part of the training process we classified against these and requested training data for anomalies with reasonable matches to items that were not expected at the site. We also requested training data for anomalies whose polarizabilities appeared to correspond to symmetric targets which did not match any of the library TOI. In all, training data were requested for twenty anomalies, three of which proved to be TOI. This phase of the demonstration took approximately 32 hours.

8.1.5 Classification

The classification algorithm itself take about 2 seconds to run. The 0.5 hr entered in Table 2 includes time for checking and re-formatting the dig list.

9.0 REFERENCES

1. Thomas Bell, "Magnetic Surface Modes and UXO/Clutter Classification and Discrimination," Final Report SERDP Project MR-1658, December 2011.
2. Thomas Bell, "Tensor Invariant Processing for Munitions/Clutter Classification," Final Report SERDP Project MR-2100, November 2013
3. James B. Kingdon, Bruce J. Barrow, Thomas H. Bell, David C. George, Glenn R. Harbaugh and Daniel A. Steinhurst, "TEMTADS Adjunct Sensor Systems: Hand-held EMI Sensor for Cued UXO Discrimination (ESTCP MR-200807) and Man-Portable EMI Array for UXO Detection and Discrimination (ESTCP MR-200909) Final Report," Naval Research Laboratory Memorandum Report NRL/MR/6110--12-9401, April 5, 2012.
4. "ESTCP Munitions Response Live Site Demonstrations, Former Southwestern Proving Ground, Arkansas," ESTCP Demonstration Plan, August 2012.
5. ESTCP Project MR-201231 Demonstration Plan, "ESTCP Munitions Response Live Site Demonstrations, Former Southwestern Proving Ground, Arkansas," Principal Investigator, Ryan Steigerwalt, Weston Solutions, Inc., April 2013.
6. Anne Andrews and Herb Nelson, "Implementing Advanced Classification on Munitions Response Sites: A Guide to Informed Decision Making For Project Managers, Regulators, and Contractors," ESTCP Final Report, December 2011.
7. Stephen Billings, Laurens Beran and Doug Oldenburg, "Robust Statistics and Regularization for feature Extraction and UXO Discrimination," Final Report, SERDP Project MR-1629, July 2011.
8. Dean Keiswetter and Laurens Beran, "Decision Metrics," ESTCP Classification Study Demonstrators Meeting, February 9, 2015.
9. Dean Keiswetter, "2009 ESTCP UXO Classification Study, San Luis Obispo, CA," Final Report, ESTCP Project MR-200910, April 2012.
10. Herb Nelson, Katherine Kaye and Anne Andrews, "Geophysical System Verification (GSV): A Physics-Based Alternative to Geophysical Prove-Outs for Munitions Response," ESTCP Final Report, July 2009.
11. Jonathan T. Miller, Dean Keiswetter, Jim Kingdon, Tom Furuya, Bruce Barrow and Tom Bell, "Source Separation using Sparse-Solution Linear Solvers," Detection and Sensing of

Mines, Explosive Objects, and Obscured Targets XV, edited by Russell S. Harmon, John H. Holloway Jr. and J. Thomas Broach, Proc. of SPIE Vol. 7664, 2010, 766409-1-8.

APPENDIX A. HEALTH AND SAFETY PLAN

The Health and Safety Plan for the on-site activities at Southwestern Proving Ground is included in the Weston Solutions, Inc. Demonstration Plan [5].

APPENDIX B. POINTS OF CONTACT

Point of Contact (Organization)	Mailing Address	Phone and e-mail	Role in Project
Herb Nelson (ESTCP)	4800 Mark Center Drive, Suite 17D08 Alexandria, VA 22350	Office: 571-372-6400 herb.nelson@nrl.navy.mil	MR Program Manager
Ryan Steigerwalt (Weston Solutions)	1400 Weston Way West Chester, PA, 19380	Office: 267-258-2672 Ryan.Steigerwalt@WestonSolutions.com	Principal Investigator
Brian Junck (Weston Solutions)	1400 Weston Way West Chester, PA, 19380	Office: 610-701-3926 Brian.Junck@WestonSolutions.com	Co-PI
Tom Bell (Leidos)	4001 N. Fairfax Dr., Suite 7132 Arlington, VA 22203	Office: (703) 312-6288 Cell: (301) 712-7021 bellth@leidos.com	Principal Investigator