2009 ESTCP UXO CLASSIFICATION STUDY, SAN LUIS OBISPO, CA

Former Camp San Luis Obispo

March 2009

Draft 2.1
November 4, 2008

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 2

2.0 Technology 4

2.1 Geophysical Data Collection 4

2.2 Data Analysis 5

3.0 Performance Objectives 6

3.1 Objective: Maximize correct classification of non-munitions 7

3.2 Objective: Maximize correct classification of munitions 7

3.3 Objective: Specification of no-dig threshold 8

3.4 Objective: Minimize number of anomalies that cannot be analyzed 8

3.5 Objective: Correct estimation of target parameters 9

4.0 Site Description 10

4.1 Site Selection 10

4.2 Site History 10

4.3 Site Topography and Geology 11

4.4 Munitions Contamination 11

4.5 Site Geodetic Control Information 12

4.6 Site Configuration 12

5.0 Test Design 13
5.1 Conceptual Experimental Design 13
5.2 Pre-Demonstration Activities 13
5.3 Calibration Activities 17
5.4 Data Collection Procedures 17
5.5 Anomaly Selection 18
5.6 Validation 19
6.0 Data Analysis Plan 20
6.1 Preprocessing 20
6.2 Parameter Estimation 20
6.3 Training 20
6.4 Training Memo 21
6.5 Classification 21
6.6 Other Data Products 22
6.7 Performance Assessment 22
7.0 Cost Assessment 24
7.1 Cost Tracking 24
7.2 Approaches to be Compared 24
8.0 Schedule of Activities 26
9.0 References 27
Appendix A: Points of Contact A-1

Figures

Figure 4-1. Final layout of the demonstration site 12

Figure 5-1. Results of a magnetometer transect survey of a part of MRS 05 at the former Camp San Luis Obispo 15

Figure 5-2. Interpolated magnetic anomaly density map of the survey area 15

Figure 5-3. Calculated EM61-MK2 cart signal expected from a 60-mm mortar 18

Figure 6-1. Format for the prioritized anomaly list that will be submitted by each classification demonstrator 21

Figure 6-2. Example ROC curve 23

Tables

Table 3-1. Performance Objectives for This Demonstration 6

Table 4-1. Geodetic Control at the former Camp San Luis Obispo site 12

Table 7-1. Details of the costs that will be tracked by each demonstrator 24

Table 8-1. Planned Schedule of Activities for the Demonstration 26

Acronyms

11X	Depth corresponding to 11 times an object's diameter
BUD	Berkeley UXO Discriminator
CNG	California National Guard
DGM	Digital Geophysical Mapping
DSB	Defense Science Board
EE/CA	Engineering Evaluation/Cost Analysis
EMI	Electromagnetic Induction
ESTCP	Environmental Security Technology Certification Program
FAR	False Alarm Rate
FUDS	Formerly Used Defense Site
GPS	Global Positioning System
HRR	Historical Records Review
IDA	Institute for Defense Analyses
MRS	Munitions Response Site
P_{class}	Probability of Correct Classification
P_{fa}	Probability of False Alarm
QA	Quality Assurance
QC	Quality Control
ROC	Receiver Operating Characteristic
SERDP	Strategic Environmental Research and Development Program
SI	Site Investigation
UXO	Unexploded Ordnance

1.0 INTRODUCTION

1.1 BACKGROUND

In 2003, the Defense Science Board observed: "The … problem is that instruments that can detect the buried UXOs also detect numerous scrap metal objects and other artifacts, which leads to an enormous amount of expensive digging. Typically 100 holes may be dug before a real UXO is unearthed! The Task Force assessment is that much of this wasteful digging can be eliminated by the use of more advanced technology instruments that exploit modern digital processing and advanced multi-mode sensors to achieve an improved level of discrimination of scrap from UXOs [1]."

Significant progress has been made in classification technology over the past several years. To date however, testing of these approaches has been primarily limited to test sites with only limited application at live sites. Acceptance of these classification technologies requires demonstration of system capabilities at real UXO sites under real world conditions. Any attempt to declare detected anomalies to be harmless and requiring no further investigation will require demonstration to regulators of not only individual technologies, but an entire decision making process.

The FY06 Defense Appropriation contained funding for the "Development of Advanced, Sophisticated, Discrimination Technologies for UXO Cleanup" in the Environmental Security Technology Certification Program (ESTCP). ESTCP responded by conducting a UXO Classification Study at the former Camp Sibert, AL [2]. The results of this first demonstration were very encouraging. Although conditions were favorable at this site, a single target-of-interest (4.2-in mortar) and benign topography and geology, all of the classification approaches demonstrated were able to correctly identify a sizable fraction of the anomalies as arising from non-hazardous items that could be safely left in the ground. Of particular note, the contractor EM61-MK2 cart survey with analysis using commercially-available methods correctly identified more than half the targets as non-hazardous.

To build upon the success of the first phase of this study, ESTCP is sponsoring a second study in 2008 at a site with more challenging topography and a wider mix of targets-of-interest. A range at the former Camp San Luis Obispo, CA has been identified for this demonstration. This document describes the planned demonstration at San Luis Obispo.

1.2 OBJECTIVE OF THE DEMONSTRATION

There are two primary objectives of this study:

1. Test and validate detection and classification capabilities of currently available and emerging technologies on real sites under operational conditions.
2. Investigate in cooperation with regulators and program managers how classification technologies can be implemented in cleanup operations.

Within each of these two overarching objectives, there are several sub-objectives.

Technical objectives of the Study

- Test and evaluate capabilities by demonstrating and evaluating individual sensor and classification technologies and processes that combine these technologies. Compare advanced methods to existing practices and validate the pilot technologies for the following:
  - Detection of UXOs
  - Identification of features that distinguish scrap and other clutter from UXO
  - Reduction of false alarms (items that could be safely left in the ground that are incorrectly classified as UXO) while maintaining Pds acceptable to all
  - Ability to identify sources of uncertainty in the classification process and to quantify their impact to support decision making, including issues such as impact of data quality due to how data is collected
  - Quantify the overall impact on risk arising from the ability to clear more land more quickly for the same investment.
  - Include the issues of a dig-no dig decision process and related QA/QC issues
- Understand the applicability and limitations of the pilot technologies in the context of project objectives, site characteristics, suspected ordnance contamination
- Collect high-quality, well documented data to support the next generation of signal processing research

1.3 REGULATORY DRIVERS

ESTCP has assembled an Advisory Group to address the regulatory, programmatic and stakeholder acceptance issues associated with the implementation of classification in the MR process.

Objective of Advisory Group

- Help the Program Office explore a UXO classification process that will be useful to regulators and managers in making decisions.
  - Under what conditions would you consider classification?
  - What does a pilot project need to demonstrate for the community to consider not digging every anomaly as a viable alternative?
    - Methodology
    - Transparency
    - QA/QC requirements
    - Validation

- For implementation beyond the pilot project,
  - How should proposals to implement classification be evaluated?
    - Site suitability
      - Geology
      - Anomaly density
      - Site topography
      - Level of understanding of expected UXO types
    - Track record on like sites
    - Performance on test site or small subset of site
    - Understanding and management of uncertainties
  - Define data needs to support decisions, particularly with regard to decisions not to dig all detected anomalies
  - Define acceptable end-products to support classification decisions

- In support of the above, provide input and guidance to the Program Office
  - Pilot project objectives and flow-down to metrics
  - Flow down of program objectives to data quality objectives
  - Demonstration/Data collection plans
  - QA/QC requirements and documentation
  - Interpretation, Analysis, and Validation
  - Process flow for classification-based removal actions

2.0 TECHNOLOGY

This demonstration will consist of data collection using a variety of geophysical sensor systems and analysis by several groups using a number of feature extraction and classification algorithms. Details of each technology to be demonstrated will be found in the individual demonstrator work plans which will be available later in the process. A brief description of the major components of the demonstration is provided below.

2.1 GEOPHYSICAL DATA COLLECTION

The two geophysical sensors to be employed in this demonstration are magnetometers and Electromagnetic Induction (EMI) sensors.

2.1.1 Magnetometer Sensor

Magnetometers sense the local perturbation to the earth's magnetic field due to nearby ferrous metal objects. Magnetometer survey data can be analyzed to obtain the location, depth and rough size of the ferrous target. For targets more than an object length or two away from the sensors, there is little shape information in the magnetometer response so these sensors are often more useful for detection than classification. The estimated depth from analysis of magnetometer data is very accurate in most cases and can be used to constrain the analysis of EMI data and improve the ultimate classification results.

A towed array of eight magnetometers with sensor location provided by cm-level GPS will be used to acquire magnetometer data for this demonstration.

2.1.2 Electromagnetic Induction Sensors

A number of commercial and developmental EMI systems will be demonstrated. The Geonics EM61-MK2 sensor, the most widely used EMI sensor for UXO surveys, will be used in three deployment techniques, one single-sensor cart configuration and two arrays. The single-sensor cart and one of the array surveys will be performed by a UXO services firm using their existing commercial equipment. The other array will be a well-instrumented array with careful measurement of array position and orientation.

Two developmental EMI systems will be used to perform cued interrogation of a number of the anomalies detected. The Berkeley UXO Discriminator, BUD, which performed so well at Camp Sibert, will again participate. The Naval Research Lab Cued EMI array will also be used for cued interrogation of anomalies.

2.1.3 Dual-Mode Sensor

The final system to be demonstrated will be a concurrent, magnetometer/EM61-MK2 dual mode sensor. This system operates by interleaving the magnetometer and EMI measurement periods to

avoid interference between the sensors. This system will be operated by a commercial survey firm with training and guidance provided by the system developer.

2.2 DATA ANALYSIS

Data analysis for this demonstration will consist of the use of physics-based models to extract target parameters followed by the use of classification algorithms to produce a prioritized dig list. In some cases these two steps will be carried out by one demonstrator, in others, different demonstrators will perform the two, distinct analysis tasks.

2.2.1 Parameter Estimation

All of the processing approaches in the demonstration will be based on a dipole model. Other processing schemes based on more sophisticated models may be investigated in related research projects, primarily through SERDP. Each selected anomaly will be analyzed according to the processing demonstration plan for each algorithm used. Both intrinsic (size, shape, materials properties) and extrinsic (location, depth, orientation) parameters will be estimated in these analyses and a list of the relevant target parameters from each analysis compiled.

2.2.2 Classification

Several types of classification processing will be evaluated in the classification study. These include:

- Statistical classification algorithms and
- Formal rule-based classification.

There are several implementations for each of these general classifier categories (i.e., mag only data, EM61 and mag data, using polarizability as a basis.) Details of the approach or approaches to be used and the data required to implement the approach will be provided in the demonstration plan provided by each signal processing group.

3.0 PERFORMANCE OBJECTIVES

The performance objectives for this demonstration are summarized in Table 3-1. Since this is a classification demonstration, the performance objectives focus target analysis and classification; we assume that the anomalies from all targets of interest have been detected and included on the target list the analysis demonstrators will work from.

The first three objectives refer to the classification part of the demonstration with the first two referring to the best results from each approach in a retrospective analysis and the third addressing how well each demonstrator is able to specify the correct threshold in advance. The final two objectives refer to the feature extraction part of the demonstration.

Table 3-1. Performance Objectives for This Demonstration

Performance Objective	Metric	Data Required	Success Criteria
Quantitative Performance Objectives			
Maximize correct classification of munitions	Number of targets-of-interest retained.	• Prioritized anomaly lists • Scoring reports from IDA	Approach correctly classifies all targets-of-interest
Maximize correct classification of non-munitions	Number of false alarms eliminated.	• Prioritized anomaly lists • Scoring reports from IDA	Reduction of false alarms by > 30% while retaining all targets of interest
Specification of no-dig threshold	P_{class} and N_{fa} at demonstrator operating point.	• Demonstrator - specified threshold • Scoring reports from IDA	Threshold specified by the demonstrator to achieve criteria above
Minimize number of anomalies that cannot be analyzed	Number of anomalies that must be classified as "Unable to Analyze."	• Demonstrator target parameters	Reliable target parameters can be estimated for > 90% of anomalies on each sensor's detection list.
Correct estimation of target parameters	Accuracy of estimated target parameters.	• Demonstrator target parameters • Results of intrusive investigation	βs ± 20% X, Y < 15 cm (1σ) Z < 10 cm (1σ) size ± 20%

3.1 OBJECTIVE: MAXIMIZE CORRECT CLASSIFICATION OF MUNITIONS

This is one of the two primary measures of the effectiveness of this approach. By collecting high-quality data and analyzing those data with advanced parameter estimation and classification algorithms we expect to be able to classify the targets with high efficiency. This objective concerns the component of the classification problem that involves correct classification of items-of-interest..

3.1.1 Metric

The metric for this objective is the number of items on the master anomaly list that can be correctly classified as munitions by each classification approach.

3.1.2 Data Requirements

Each demonstrator will prepare a prioritized dig list for the targets on the master anomaly list. IDA personnel will use their scoring algorithms to assess the results.

3.1.3 Success Criteria

The objective will be considered to be met if all of the items-of-interest are correctly labeled as munitions on the prioritized anomaly list.

3.2 OBJECTIVE: MAXIMIZE CORRECT CLASSIFICATION OF NON-MUNITIONS

This is the second of the two primary measures of the effectiveness of this approach. By collecting high-quality data and analyzing those data with advanced parameter estimation and classification algorithms we expect to be able to classify the targets with high efficiency. This objective concerns the component of the classification problem that involves false alarm reduction.

3.2.1 Metric

The metric for this objective is the number of items-of-interest on the master dig list that can be correctly classified as non-munitions by each classification approach.

3.2.2 Data Requirements

Each demonstrator will prepare a prioritized dig list for the targets on the master anomaly list. IDA personnel will use their scoring algorithms to assess the results.

3.2.3 Success Criteria

The objective will be considered to be met if more than 30% of the non-munitions items can be correctly labeled as non-munitions while retaining all of the targets-of-interest on the dig list.

3.3 OBJECTIVE: SPECIFICATION OF NO-DIG THRESHOLD

In a retrospective analysis as will be performed in this demonstration, it is possible to tell the true classification capabilities of a classification procedure based solely on the prioritized dig list submitted by each demonstrator. In a real-world scenario, all targets may not be dug so the success of the approach will depend on the ability of an analyst to accurately specify their dig/no-dig threshold.

3.3.1 Metric

P_{class} and number of false alarms, N_{fa}, at the demonstrator-specified threshold are the metrics for this objective.

3.3.2 Data Requirements

Each demonstrator will prepare a ranked anomaly list with a dig/no-dig threshold indicated. IDA personnel will use their scoring algorithms to assess the results.

3.3.3 Success Criteria

The objective will be considered to be met if more than 30% of the non-munitions items can be correctly labeled as non-munitions while retaining all of the targets-of-interest at the demonstrator-specified threshold.

3.4 OBJECTIVE: MINIMIZE NUMBER OF ANOMALIES THAT CANNOT BE ANALYZED

Anomalies for which reliable parameters cannot be estimated cannot be classified by the classifier. These anomalies must be placed in the dig category and reduce the effectiveness of the classification process.

3.4.1 Metric

The number of anomalies for which reliable parameters cannot be estimated is the metric for this objective.

3.4.2 Data Requirements

Each demonstrator that estimates target parameters will provide a list of all parameters as part of their results submission along with a list of those anomalies for which parameters coiuld not be reliable estimated.

3.4.3 Success Criteria

The objective will be considered to be met if reliable parameters can be estimated for > 90% of the anomalies on each sensor anomaly list.

3.5 OBJECTIVE: CORRECT ESTIMATION OF TARGET PARAMETERS

This objective is intended involves the accuracy of the target parameters that are estimated in the first phase of the analysis. Successful classification is only possible if the input features are internally consistent. The obvious way to satisfy this condition is to estimate the various target parameters accurately.

3.5.1 Metric

Accuracy of estimation of target parameters is the metric for this objective.

3.5.2 Data Requirements

Each demonstrator that estimates target parameters will provide a list of all parameters as part of their results submission. IDA analysts will compare these estimated parameters to those measured during the intrusive investigation and determined via subsequent in-air measurements.

3.5.3 Success Criteria

The objective will be considered to be met if the estimated βs are within ± 20%, the estimated X, Y locations are within 15 cm (1σ), the estimated depths are within 10 cm (1σ), and the estimated size is within ± 20%.

4.0 SITE DESCRIPTION

The site description material reproduced is here is taken from the recent SI report [3]. More details can be obtained in the report. The former Camp San Luis Obispo is approximately 2,101 acres situated along Highway 1, approximately five miles northwest of San Luis Obispo, California. The majority of the area consists of mountains and canyons. The site for this demonstration is a mortar target on hilltop in Munitions Response Site (MRS) 05 (within former Rifle Range #12).

4.1 SITE SELECTION

This site was chosen as the next in a progression of increasingly more complex sites for demonstration of the classification process. The first site in the series, Camp Sibert, had only one target-of-interest and item "size" was an effective discriminant. At this site, there are at least three targets-of-interest: 60-mm, 81-mm, and 4.2-in mortars. This introduces another layer of complexity into the process.

4.2 SITE HISTORY

Camp San Luis Obispo was established in 1928 by California as a National Guard Camp. Identified at that time as Camp Merriam, it originally consisted of 5,800 acres. Additional lands were added in the early 1940s until the acreage totaled 14,959. During World War II, Camp San Luis Obispo was used by the U.S. Army from 1943 to 1946 for infantry division training including included artillery, small arms ranges, mortar, rocket, and grenade ranges. According to the Preliminary Historical Records Review (HRR), there was a total of 27 ranges and thirteen training areas located on Camp San Luis Obispo during World War II. Construction at the camp included typical dwellings, garages, latrines, target houses, repair shops, and miscellaneous range structures. Following the end of World War II, a small portion of the former camp land was returned to its former private owners. The U.S. Army was making arrangements to relinquish the rest of Camp San Luis Obispo to the State of California and other government agencies when the conflict in Korea started in 1950. The camp was reactivated at that time.

The U.S. Army used the former camp during the Korean War from 1951 through 1953 where the Southwest Signal Center was established for the purpose of signal corps training. The HRR identified eighteen ranges and sixteen training areas present at Camp San Luis Obispo during the Korean War. A limited number of these ranges and training areas were used previously during World War II. Following the Korean War, the camp was maintained in inactive status until it was relinquished by the Army in the 1960s and 1970s. Approximately 4,685 acres was relinquished to the General Services Administration (GSA) in 1965. GSA then transferred the property to other agencies and individuals beginning in the late-1960s through the 1980s; most of which was transferred for educational purposes (Cal Poly and Cuesta College). A large portion of Camp San Luis Obispo (the original 5,880 acres) has been retained by the California National Guard (CNG) and is not part of the FUDS program

4.3 SITE TOPOGRAPHY AND GEOLOGY

The Camp San Luis Obispo site consists mainly of mountains and canyons classified as grassland, wooded grassland, woodland, or brush. A major portion of the site is identified as grassland and is used primarily for grazing. Los Padres National Forest (woodland) is located to the north-northeastern portion of the site. During the hot and dry summer and fall months, the intermittent areas of brush occurring throughout the site become a critical fire hazard.

The underlying bedrock within the Camp San Luis Obispo site area is intensely folded, fractured, and faulted. The site is underlain by a mixture of metamorphic, igneous, and sedimentary rocks less than 200 million years old. Scattered throughout the site are areas of fluvial sediments overlaying metamorphosed material known as Franciscan mélange. These areas are intruded by plugs of volcanic material that comprise a chain of former volcanoes extending from the southwest portion of the site to the coast. Due to its proximity to the tectonic interaction of the North American and Pacific crustal plates, the area is seismically active.

A large portion of the site consists of hills and mountains with three categories of soils occurring within: alluvial plains and fans; terrace soils; and hill/mountain soils. Occurring mainly adjacent to stream channels are the soils associated with the alluvial plains and fans. Slope is nearly level to moderately sloping and the elevation ranges from 600 to 1,500 feet. The soils are very deep and poorly drained to somewhat excessively drained. Surface layers range from silty clay to loamy sand. The terrace soils are nearly level to very steep and the elevations ranges from 600 to 1,600 feet. Soils in this unit are considered shallow to very deep, well drained, and moderately well drained. The surface layer is coarse sandy loam to shaley loam. The hill/mountain soils are strongly sloping to very steep. The elevation ranges from 600 to 3,400 feet. The soils are shallow to deep and excessively drained to well drained with a surface layer of loamy sand to silty clay.

4.4 MUNITIONS CONTAMINATION

A large variety of munitions have been reported as used at the former Camp San Luis Obispo. Munitions debris from the following sources was observed in MRS 05 during the 2007 SI:

- 4.2-inch white phosphorus mortar
- 4.2-inch base plate
- 3.5-inch rocket
- 37mm
- 75mm
- 105mm
- 60mm mortar
- 81mm mortar

- practice bomb
- 30 cal casings and fuzes.
- flares found of newer metal; suspected from CNG activities

At the particular site of this demonstration, 60-mm, 81-mm, and 4.2-in mortars and mortar fragments have been observed. The excavation of two grids as part of the preparatory activities will provide information on the munitions at this target site.

4.5 SITE GEODETIC CONTROL INFORMATION

Table 4-1. Geodetic Control at the former Camp San Luis Obispo site

ID	Latitude	Longitude	Elevation (m)	Northing (m)	Easting (m)	HAE (m)
ESTCP	35º 20' 37.77465" N	120º 44' 25.95073"W	113.69	3,913,515.95	705,330.89	76.01

4.6 SITE CONFIGURATION

The demonstration site will be configured as one 11.8-acre area. It will span the hillside that is the historical mortar target. The cart systems will survey 45 30- x 30-m grids within this area for a total of 10 acres. The vehicular systems will survey the entire area. Details of the final site extent are shown in Figure 4-1. The calibration strip and training pit will be located off the site, convenient to the access road.

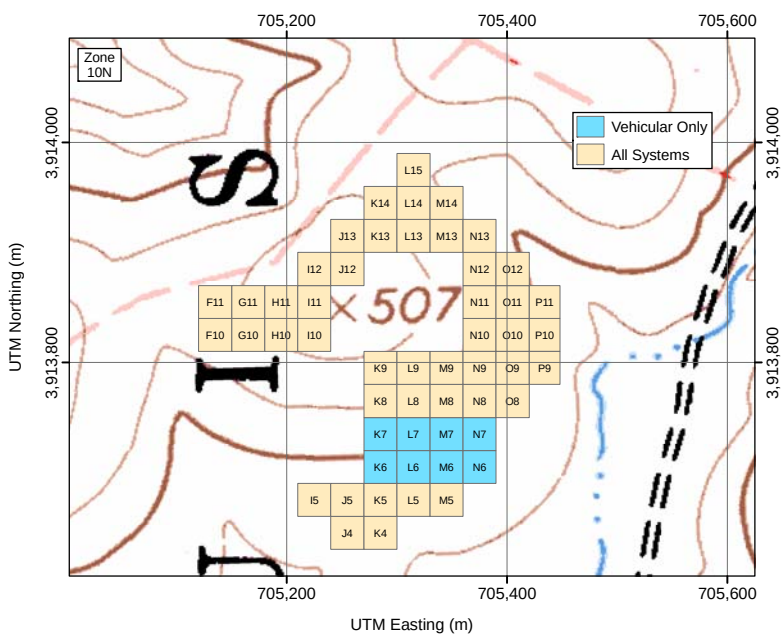

Figure 4-1. Final layout of the demonstration site showing the grids to be surveyed by all systems (10 acres) and the additional 8 grids (1.8 acres) to be surveyed by the vehicular systems.

5.0 TEST DESIGN

5.1 CONCEPTUAL EXPERIMENTAL DESIGN

The objective of this program is to demonstrate a methodology for the use of classification in the munitions response process. The three key components of this methodology are collection of high-quality geophysical data and principled selection of anomalous regions in those data, analysis of the selected anomalies using physics-based models to extract target parameters such as size, shape, and materials properties, and the use of those parameters to construct a prioritized dig list. Each of these components will be handled separately in this program.

The ESTCP Program Office will coordinate data collection activities. This will include all preparatory activities, arranging for a data collection by well-validated systems, selection of anomalies for analysis from each geophysical data set, and compilation of the individual sensor anomaly lists into a master list.

Data analysis demonstrators will then process the individual data sets using existing routines to extract target parameters. These parameters are passed to the classification routines which, after training on a limited amount of site-specific ground truth, are used to produce prioritized dig lists.

Validation digging will be coordinated by the Program Office. Since this is a demonstration, all anomalies on the master dig list will be investigated. The underlying target will be uncovered, photographed, located with a cm-level GPS system, and removed. The identities of a small number of the recovered items will be provided to the demonstrators as training data; they will use these inputs to finalize algorithms and adjust thresholds.

At the conclusion of training, each demonstrator will submit a prioritized dig list for each data set they have analyzed. These lists will be ordered from the item the demonstrator is most confident is not hazardous through the item the demonstrator is most confident is a munition. The anomalies for which the demonstrator was not able to extract meaningful parameters will be placed at the bottom of the list. These inputs will be scored by the Institute for Defense Analyses with emphasis on the number of items that are correctly labeled non-hazardous while correctly labeling all munitions items.

The primary objective of the demonstration will be to assess how well each demonstrator is able to order their prioritized anomaly list and specify the threshold separating high confidence clutter from all other items. The secondary objective will be to determine the classification performance that could be achieved by each approach through a retrospective analysis.

5.2 PRE-DEMONSTRATION ACTIVITIES

Pre-demonstration activities include:

- Magnetometer transects to assist in choice of study site.

- Collection of historical records about the site.

- Identification or establishment of first order navigation points to be used for all emplacement, data collection and validation activities.

- EMI survey of approximately 30 acres, to be used to guide selection of the 10 acre demonstration site, calibration strip area, and the characterization grids as well as to guide the emplacement of seed targets (discussed below). To preserve the integrity of a blind demonstration, this survey was conducted by personnel not involved in the later data collection and analysis.

- Based on the EMI survey, two 50' x 50' grids will be dug, in the manner generally used in EE/CAs. The dig results will provide site-specific information to guide the selection of targets-of-interest for the site, establish the depth distributions required for the seed items, and be available for use by the demonstrators.

- Develop a seed plan and seed the site.

- Establish a calibration strip and training pit near the demonstration area.

5.2.1 Initial Magnetometer Transects

As part of the site selection process, the ESTCP Program Office conducted limited magnetometer transect surveys of two candidate sites. The results from the San Luis Obispo site are shown in Figures 5-1 and 5-2. Figure 5-1 shows the actual transect course over ground and the positions of the detected anomalies. Figure 5-2 shows the interpolated magnetometer anomaly densities across the site calculated from the survey data. It is clear from both figures that there is a range of anomaly densities across the site with two high-density areas on the hillside.

5.2.2 Survey of Historical Records

Much of the historical information on this site has been collected in the SI report [3]. This report is posted on the ESTCP ftp server and can be used for reference.

5.2.3 First-order Navigation Point

To avoid confusion between and among various demonstrators, it is important that all survey data and validation activities be conducted on a common coordinate system. On August 13, 2008, Cannon Associates, a licensed survey firm in San Luis Obispo, installed a permanent monument labeled “ESTCP” the coordinates of which are given in Table 4-1.

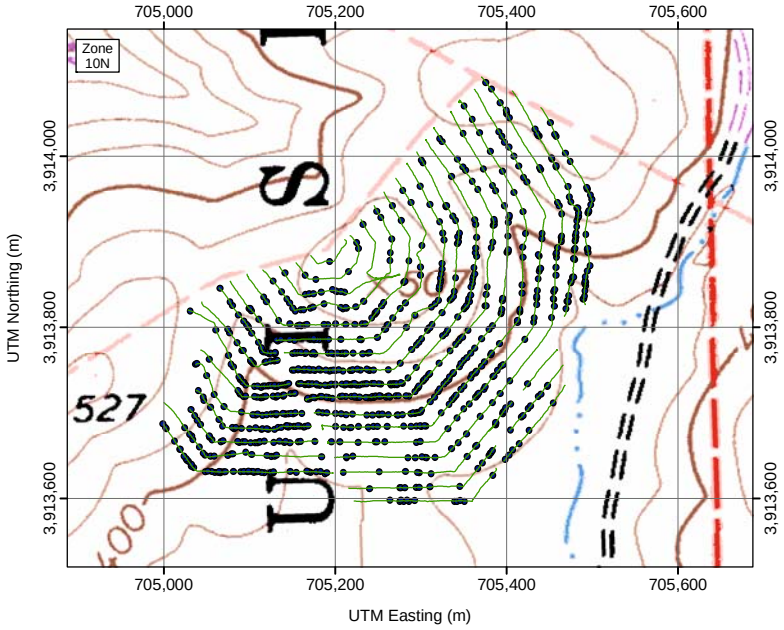

Figure 5-1. Results of a magnetometer transect survey of a part of MRS 05 at the former Camp San Luis Obispo. The lines are the actual transect course-over-ground and the points are the location of detected anomalies.

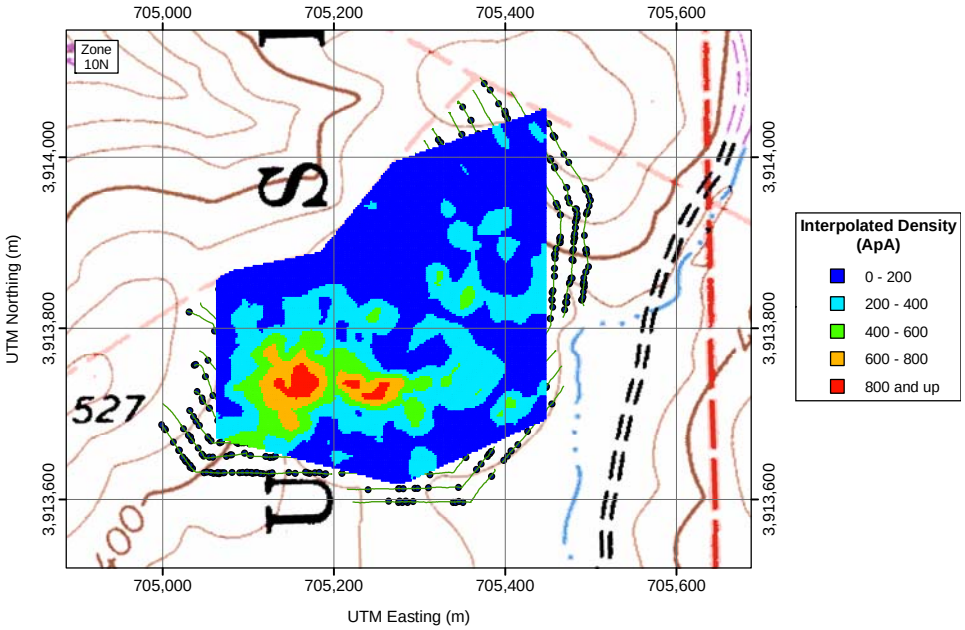

Figure 5-2. Interpolated magnetic anomaly density map of the survey area shown in Figure 5-1.

5.2.4 Initial EMI Survey

The details of individual anomaly locations within the 30-acre demonstration site are required for successful conduct of this demonstration. As in the first demonstration in this series, the 10 acre demonstration area will be chosen from a portion of the site where the anomaly density is low enough to guarantee that a large majority of the anomalies are isolated from each other. Data processing techniques that can reliably extract target parameters from overlapping anomaly signatures have not been well validated to date. The demonstration area will be chosen so that a limited number of overlapping anomalies are in the survey data. They will be left on the anomaly list with the presumption that most analysis methods will declare them as anomalies for which reliable parameters cannot be extracted.

Based on the survey results, two grids from which to develop site-specific target and clutter information were chosen from the higher-density parts of the site and a tentative survey area defined (Figure 4-1). These initial survey data will be used to guide the development of the seed plan.

5.2.5 Acquire Site-specific Information

Although the historical records are not definitive, it is thought that the targets of interest on this part of former Camp San Luis Obispo are 60-mm, 81-mm, and 4.2-in mortars. To confirm these targets-of-interest, to get an indication of their depth profile on this site, and to provide to the demonstrators some information about the clutter environment at the site, two 50' x 50' grids will be completely excavated.

The initial dig list for this investigation will be developed from the EMI survey described above. Each target will be uncovered, photographed, and its depth recorded. From the initial density estimates presented in Figure 5-2, we expect to encounter 50 to 75 targets in each grid. After all targets on the list have been removed, the UXO specialist will perform a “Mag and Flag” survey of each grid to ensure that all targets have been removed.

5.2.6 Seed the Site

At a live site such as this, the ratio of clutter to targets-of-interest is such that only a small number of targets-of-interest may be found in a 10- 15 acre area; far from enough to determine any demonstrator’s classification performance with acceptable confidence bounds. To avoid this problem, the site will be seeded with enough targets-of-interest to ensure reasonable statistics.

The details of the seed plan will be based on the results of the excavation of the initial grids which will be chosen in high-density areas to maximize the number of targets recovered. These digs will confirm the identity of the targets-of-interest and, if a sufficient number of intact items-of-interest are recovered, provide a guide to the depth distribution of targets-of-interest on this site. If the depth distributions for the targets-of-interest cannot be confidently determined, seed items will be buried over depths from just covered to 11x the items diameter if the location of the

bedrock layer allows. As in the case of the Sibert study, the depths of burial will be biased toward the shallower half of the distribution.

Although it is reasonable to expect to find more *in situ* targets-of-interest on this site, the seed plan will be developed using enough of each target-of-interest to obtain reasonable statistics in the absence of any *in situ* targets-of-interest. To the extent possible, items recovered from other live ranges will be used as seeds. If availability precludes this, items from the inventory will be thoroughly degaussed prior to emplacement.

5.2.7 Establish a Calibration Strip and Training Pit

A quiet area near the entrance to the demonstration area will be located to establish a sensor calibration strip to be used for daily verification of proper sensor operation and a training pit to be used to collect sensor data for algorithm training. Each of these items will be discussed in more detail in later sections.

5.3 CALIBRATION ACTIVITIES

A calibration strip and training pit will be provided for calibration purposes. The calibration strip is intended for daily function check of the survey equipment. It will consist of a small number of targets-of-interest and standard objects (such as steel spheres) buried at known depths and locations. Each demonstrator will survey the strip each morning and evening of survey work. By comparing the twice-daily responses to the expected responses from these targets, the demonstrator can verify that her equipment is operating properly.

All other calibration activities required will be accomplished using the training pit. An example of each of the targets-of-interest can be placed in the pit at a variety of depths and orientations and the response measured. These data can be used to set detection thresholds and provide additional training data to the classification demonstrators.

5.4 DATA COLLECTION PROCEDURES

5.4.1 Scale of Demonstration

The demonstration will be conducted at the Former Camp San Luis Obispo, CA. The 10- 15 acre demonstration area will be chosen based on the results of the initial EMI survey.

5.4.2 Sample Density

Survey data with the commercial EMI instruments will be collected with a line spacing of 0.5-m. One of the EM61-MK2 arrays will collect data in two orthogonal directions doubling the sample density. The magnetometer data will be collected with 0.25-m line spacing.

5.4.3 Quality Checks

Details of any data quality checks that will be performed in addition to the twice-daily survey of the calibration strip will be found in the individual demonstrator demonstration plans.

5.4.4 Data Handling

Each data collection demonstrator will provide raw and located, preprocessed data to the ESTCP Program Office for archiving. In consultation with the Program Office, each demonstrator will identify anomalies in their data set using procedures outlined in the next section.

5.5 ANOMALY SELECTION

Anomalies will be selected from each geophysical data set using a target response-based threshold. An example of the response of an EM61-MK2 cart to a 60-mm mortar as a function of depth is shown in Figure 5-3. Plotted in this figure are the calculated responses when the mortar is in its most favorable and least favorable orientation and a number of survey measurements from a training pit that validate the predictions. From this plot, we can predict the minimum signal expected from this sensor for this target at any depth. Once the maximum depth-of-interest is established, a detection threshold can be established by adjusting the minimum signal downward by a safety factor (50% for this program).

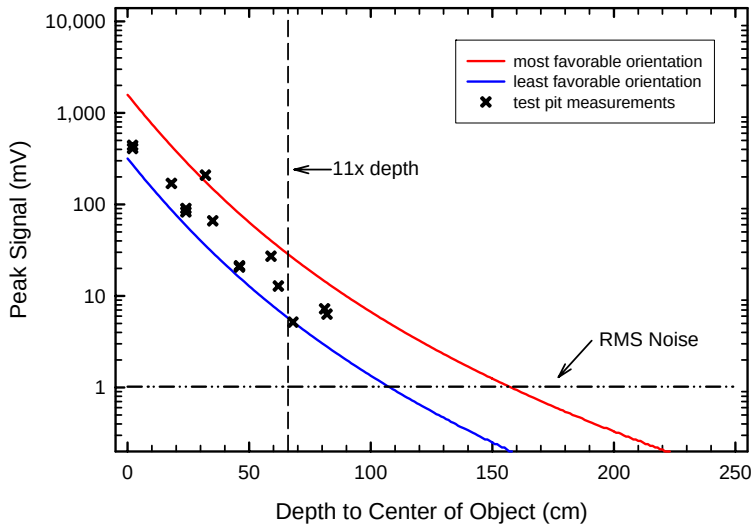

Figure 5-3. Calculated EM61-MK2 cart signal expected from a 60-mm mortar when the mortar is in its most favorable and least favorable orientation. Actual survey measurements from a training pit are shown as validation for the predictions. The depth corresponding to 11x the mortar's diameter is marked.

The depth-of-interest for each item will be set at the depth of the rock layer or 11x the items diameter, whichever is deeper. The detection threshold for that item will then be fixed at half the minimum signal expected from that item at the depth-of-interest. Since multiple targets-of-interest are expected at this site, the ultimate anomaly selection threshold for each data set will be the smallest of the individual item thresholds.

5.6 VALIDATION

At the conclusion of data collection activities, all anomalies on the master anomaly list will be excavated. Each item encountered will be identified, photographed, its depth measured, its location determined using cm-level GPS, and the item removed if possible. All non-hazardous items will be saved for later in-air measurements as appropriate.

6.0 DATA ANALYSIS PLAN

6.1 PREPROCESSING

Survey data will be preprocessed (located and simple filtering) by the data collection demonstrators in preparation for anomaly selection. These preprocessed data will be provided to each analysis demonstrator for use in target parameter extraction. If a particular demonstrator desires more control over the preprocessing steps, the raw survey data will be held by the Program Office and provided as appropriate.

6.2 PARAMETER ESTIMATION

All of the processing approaches to be used in this demonstration are based on a dipole model. Other processing schemes based on more sophisticated models may be investigated in related research projects, primarily through SERDP.

Each selected target will be analyzed according to the processing demonstration plan for each algorithm. The analysis results will be evaluated for correctness of parameter estimation. The parameters will be fed into the signal processing approaches below and ranked dig lists will be generated. Demonstrators will be required to specify thresholds beyond which all detections are clutter. All targets from all dig lists will be dug. If resources constrain, digs will be confined to 100% of a portion of the site.

Several types of classification processing will be evaluated in the classification study. These include:

- Statistical classification algorithms
- Formal rule-based classification.

There are several implementations for each of these general classifier categories (i.e., mag only data, EM61 and mag data, using magnetic remnance as a basis, using polarizability as a basis.) Details of the approach or approaches to be used and the data required to implement the approach will be provided in the demonstration plan provided by each signal processing group.

6.3 TRAINING

All of the classification approaches will require some level of training data. These data will come from three sources:

- Sensor data for the targets-of-interest collected in previous testing,
- Data collected over the training pit, and
- Ground truth from a small fraction of the excavations provide to all demonstrators.

The ESTCP Program Office will convene a meeting of all demonstrators prior to the start of data collection. At this meeting, the various data analysis demonstrators can list the additional training data they will require from the training pit and the data collection plans can be modified accordingly. Individual analysts may request either munitions or clutter data or a mix of both.

6.4 TRAINING MEMO

After training their algorithms with the training data provided, each classification demonstrator will submit a training memo report to the Program Office. This report will detail the criteria used to assign anomalies to the "can't analyze" class, discuss the parameters used for classification and specify the values of all adjustable parameters that will used in the final classification process.

6.5 CLASSIFICATION

Following acceptance of their training memo report, each demonstrator will produce a ranked anomaly list for each of the sensor data sets they are processing. The list will follow the format shown in Figure 6-1.

Rank	Anomaly ID	$P_{clutter}$	Comment
1	247	.97	
2	1114	.96	High confidence NOT munition
3	69	…	
…	…	…	
…	…	…	
…	…	…	Can't make a decision
…	…	…	
…	…	…	
…	…	…	
…	…	…	High confidence munitions
…	…	.03	
…	…	.02	
…	…		
N-2	…		
N-1	…		Can't extract reliable features
N	…		

Threshold

Figure 6-1. Format for the prioritized anomaly list that will be submitted by each classification demonstrator.

The first item on each anomaly list will be that item which the demonstrator is the most confident is not a munition. The items will be ranked according to decreasing confidence that the item is not hazardous. The demonstrator will provide two thresholds. The first threshold, or

demonstrator operating point, is indicated in the figure and corresponds to the last item that can be classified as "high confidence not a munition." The demonstrators will also indicate which item is the first that can be classified as "high confidence munition." All targets for which reliable parameters cannot be extracted must be dug and will be placed at the bottom of the list.

Of course, it is possible that groups of anomalies in the top three categories could have equal rank. Those in the "can't extract reliable features" group can be listed in any order as there is no way to distinguish among them.

Along with the rank, each demonstrator will provide the basis for the ranking. In the example shown in Figure 6-1, the demonstrator is able to derive a probability of being clutter and uses this as the basis for the ranking. This is the preferred approach but not all classification procedures result in a true probability. In those cases, the demonstrator will supply the basis for the ranking along with a discussion of its meaning. Each of the items for which reliable features can not be extracted will be labeled by the criterion listed in the training memo that was used to make this classification.

6.6 OTHER DATA PRODUCTS

In addition to the training results White Paper and the prioritized dig list, each demonstrator that extracts target parameters will submit these also. After the prioritized lists are scored and the validation results released, each demonstrator will submit a final report that includes a failure analysis addressing, at a minimum, why any items of interest were classified as non-munitions or "can't extract reliable parameters."

6.7 PERFORMANCE ASSESSMENT

All demonstrator dig lists will be scored against the emplaced and recovered targets by the Institute for Defense Analyses (IDA). An example ROC curve is shown in Figure 6-2, with the areas of interest for the analysis indicated.

Point A in the figure corresponds to the total number of non-hazardous items in the test portion of the data set. This is the number of unnecessary digs that would be required in the traditional approach. Point B marks the demonstrator-specified threshold delineating high-confidence non-munitions from other targets. According to the classification by the demonstrator, the targets to the right of point B could be safely left in the ground or dug in a less costly manner. Point C corresponds to the maximum number of non-hazardous items when all the munitions have been correctly identified by the demonstrator. This quantity is only available from retrospective analysis. Finally, Point D corresponds to the targets for which reliable features could not be determined. In the example shown, there were a few munitions items among those targets along with nearly 150 clutter items. All of these targets must be dug as if they were hazardous.

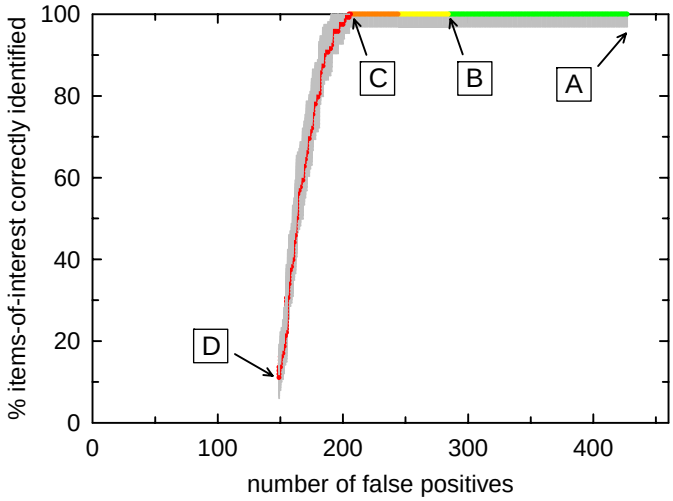

Figure 6-2. Example ROC curve with four points of interest indicated. Point A corresponds to the total number of non-munitions items in the test data set, all of which would have to be dug in the traditional approach. Point B corresponds to the demonstrator-specified threshold, Point C corresponds to the maximum possible number of avoided digs from a retrospective analysis, and Point D corresponds to items for which reliable parameters could not be extracted.

IDA will construct similar ROC curves for each anomaly list submitted by each demonstrator. In the first demonstration in this series, that amounted to over 200 ROC curves. As discussed in Section 3, the primary performance metric will be the point at which the ROC curve reaches 100% identification of targets-of-interest. The number of clutter items correctly identified at this point is a measure of the savings possible for each method. As a secondary metric, the number of items before point D, those that could not be reliably analyzed, will be assessed.

7.0 COST ASSESSMENT

7.1 COST TRACKING

Each individual technology demonstrator will report their survey costs in the format requested by ESTCP in their final reports. Overall demonstration costs will be compiled from these component costs and the costs not attributable to an individual demonstration. The costs to be tracked are detailed in Table 7-1.

Table 7-1. Details of the costs that will be tracked by each demonstrator.

Cost Category	Description
Presurvey activities	The presurvey costs for each field activity will be tied to that activity. For example, for DGM it will include the costs of planning and programming the navigation system and setting up the GPS.
Survey costs	The survey costs will include the time spent in the field collecting and recording data
Processing Costs	These costs will include the preprocessing time spent by the data collector to merge the location and geophysics data and perform standard data clean up activities and the costs of the entire data processing flow conducted by each demonstrator to select, analyze and classify targets (as appropriate)
Digging Costs	Digging costs will be estimated as the number of holes required to satisfy the dig list requirements for each processing approach times the cost per dig, as those targets would be dug with the available information

7.2 APPROACHES TO BE COMPARED

Two remediation approaches will be compared in the final report for this demonstration.

- **DGM and Dig**. This corresponds to current practice. The geophysical data are mapped and anomalies are selected based on a predetermined threshold. All targets on the anomaly list are dug with full safety precautions.

- **DGM, Characterize, Classify, and Dig**. This approach involves full implementation of post-detection signal processing. Anomalies are selected using target-of-interest based thresholds. Each anomaly is analyzed using physics-based models to extract target parameters. These parameters are submitted to a classifier which produces a ranked dig

list and a threshold is determined. Targets that are classified as high confidence non-munitions can be left in the ground or, at the discretion of site stakeholders, dug with reduced costs by eliminating some of the precautions required when digging hazardous items.

8.0 SCHEDULE OF ACTIVITIES

The planned schedule of activities for this demonstration is listed in Table 8-1.

Table 8-1. Planned Schedule of Activities for the Demonstration

Date	Event
16-Jun-08	Demonstration plan to Demonstrators and AG
20-Jun-08	EM61 presurvey to support selection of demonstration area and grids and support seeding
15-Jul-08	Draft Dem Plans from Data Collectors to All
4-Aug-08	Meet with Demonstrators
15-Oct-08	Dig Grids
5-Nov-08	Seeding and calibration lane plans
8-Dec-08	Seed and cal lane emplacement
1-Feb-09	Final demonstration plans from data collectors
1-Mar-09	Draft Demo Plans from Processors
15-Mar-09	Data Collection Begins
1-Apr-09	Final Demonstration plans from processors
31-May-09	Data Collection Complete
15-Jun-09	Target Selection/Detection
1-Jul-09	Validation Plan
15-Jul-09	Digging
15-Aug-09	Training Data to Processors
1-Sep-09	Training Memos Due from Processors
1-Oct-09	Dig Lists from Processors
15-Nov-09	Scoring Complete
1-Jan-10	Draft Reports from Processors
1-Mar-10	Draft PO Final Report V0 to AG
15-Mar-10	FINAL reports from Demonstrators
31-May-10	FINAL PO Report

9.0 REFERENCES

1. "Report of the Defense Science Board Task Force on Unexploded Ordnance," December 2003, Office of the Under Secretary of Defense for Acquisition, Technology, and Logistics, Washington, D.C. 20301-3140, http://www.acq.osd.mil/dsb/reports/uxo.pdf.

2. "ESTCP Pilot Program, Classification Approaches in Munitions Response," Nelson, H., Kaye, K., and Andrews, A.

3. "Final Site Inspection Report, Former Camp San Luis Obispo, San Luis Obispo, CA," Parsons, Inc., September 2007.

Appendix A: Points of Contact

POINT OF CONTACT	ORGANIZATION	Phone Fax e-mail	Role in Project
Dr. Jeff Marqusee	ESTCP Program Office 901 North Stuart Street, Suite 303 Arlington, VA 22203	703-696-2120 (V) 703-696-2114 (F) jeffrey.marqusee@osd.mil	Director, ESTCP
Dr. Anne Andrews	ESTCP Program Office 901 North Stuart Street, Suite 303 Arlington, VA 22203	703-696-3826 (V) 703-696-2114 (F) anne.andrews@osd.mil	Deputy Director, ESTCP
Dr. Herb Nelson	Naval Research Lab Chemistry Division Code 6110 Washington, DC 20375	202-767-3686 (V) 202-404-8119 (F) 202-215-4844 (C) herb.nelson@nrl.navy.mil	Program Manager, MM
Ms. Katherine Kaye	HydroGeoLogic, Inc. 11107 Sunset Hills Road, Suite 400 Reston, VA 20190	410-884-4447 (V) kkaye@hgl.com	Program Manager Assistant, MM
Dr. Shelley Cazares	Institute for Defense Analyses 4850 Mark Center Drive Alexandria, VA 22311	703-845-6792 (V) 703-578-2877 (F) scazares@ida.ord	Performance Assessment
demonstrators to be added as appropriate			