FINAL REPORT

2009 ESTCP UXO Discrimination Study, San Luis Obispo, CA

ESTCP Project MR-200501

NOVEMBER 2010

Lawrence Carin
Levi Kennedy
Xianyang Zhu
Nilanjan Dasgupta
Signal Innovations Group

This document has been cleared for public release

2009 ESTCP UXO DISCRIMINATION STUDY, SAN LUIS OBISPO, CA

ESTCP Project # MM-200501

Authors: Lawrence Carin (PI), Levi Kennedy, Xianyang Zhu and Nilanjan Dasgupta

Signal Innovations Group
1009 Slater Road, Suite 200
Research Triangle Park, NC 27703

Submission Date: November 3, 2010
Version 3

Contents

1 INTRODUCTION **1**
- 1.1 BACKGROUND . . . 1
- 1.2 OBJECTIVE OF THE DEMONSTRATION . . . 1
- 1.3 REGULATORY DRIVERS . . . 2

2 TECHNOLOGY - UXO DISCRIMINATION **3**
- 2.1 FORMULATION OF THE FORWARD MODEL . . . 3
- 2.2 ANALYTICAL COMPUTATION OF THE MAGNETIC FIELD . . . 5
- 2.3 SUPERVISED BAYESIAN CLASSIFICATION . . . 7
- 2.4 SEMI-SUPERVISED LEARNING . . . 8
- 2.5 ACTIVE LEARNING . . . 8
- 2.6 ADVANTAGES AND LIMITATIONS OF TECHNOLOGY . . . 9

3 PERFORMANCE OBJECTIVES **10**

4 SITE DESCRIPTION **11**
- 4.1 TEST SITE SELECTION . . . 11
- 4.2 TEST SITE HISTORY AND CHARACTERISTICS . . . 12

5 TEST DESIGN **13**

6 DATA ANALYSIS **13**
- 6.1 PROCEDURE FOR DATA CHIP SELECTION . . . 13
- 6.2 MODEL PARAMETER EXTRACTION . . . 16
- 6.3 FEATURE SELECTION . . . 17

7 PERFORMANCE ASSESSMENT **21**
- 7.1 PERFORMANCE ANALYSIS OF PASSIVE CLASSIFIERS . . . 22
  - 7.1.1 SETTING OF CLASSIFICATION THRESHOLDS . . . 23
  - 7.1.2 TEMTADS CLASSIFIER PERFORMANCE . . . 24
  - 7.1.3 METALMAPPER CLASSIFICATION PERFORMANCE . . . 25
  - 7.1.4 BUD CLASSIFICATION PERFORMANCE . . . 27
  - 7.1.5 MSEMS CLASSIFICATION PERFORMANCE . . . 27
  - 7.1.6 EM61-ARRAY CLASSIFICATION PERFORMANCE . . . 28
  - 7.1.7 EM61-CART CLASSIFICATION PERFORMANCE . . . 29
  - 7.1.8 MAG CLASSIFICATION PERFORMANCE . . . 29
- 7.2 ACTIVE LEARNING CLASSIFICATION PERFORMANCE . . . 30

8 COST ASSESSMENT **32**

9 IMPLEMENTATION ISSUES **33**
- 9.1 ENVIRONMENTAL CHECKLIST . . . 33
- 9.2 END-USER ISSUES . . . 33

List of Figures

1 The figure shows the geometry of a rectangular coil (shown in red) in the x-y plane. Our objective is to obtain an analytical expression of the magnetic field at any location P(x,y,z) due to an unit current flowing through the rectangular coil. . . 6

2 Performance objectives for the Camp San Luis Obispo study 11

3 (a) The figure shows a 5mx5m data chip centered around anomaly ID 175. The drawn polygon is visually inspected to contain only one anomaly, (b) The figure shows 5mx5m data chip centered around anomaly ID 489. The drawn polygon contains two anomalies. The data contained within the polygon is subjected to feature extraction model to obtain two sets of anomaly parameters, (c) The figure shows 5mx5m data chip centered around anomaly ID 548. Three sets of anomaly parameters are obtained jointly from the data contained within the polygon. 14

4 (a) The figure shows 5mx5m data chip centered around anomaly ID 574. The anomaly is too close to very strong signal from a neighboring anomaly, (b) The figure shows 5mx5m data chip centered around anomaly ID 782. There are too many anomalies in close proximity, making it difficult to extract their parameters reliably, (c) The figure shows 5mx5m data chip centered around anomaly ID 999. Signal-to-noise ratio is considered too low for reliable feature extraction. 15

5 (a) Three principal magnetic polarizabilities of a 81mm mortar (ID:1081) is shown as a function of the seven timegates chosen for inversion. We observe M_1 to be significantly larger than other two polarizabilities. M_2 and M_3 are nearly identical. The response is expected typically of a cylindrical anomaly (81mm mortar is a cylindrical UXO), (b) The plot shows the fitting error of 500 random initializations for anomaly ID 1081 (a 81mm mortar). We observe good fit with the measured field data for 491 initiations, and all those 491 initializations converged to the parameter vector shown in (a). 17

6 (a) Three principal magnetic polarizabilities of a 4.2-inch mortar (ID:473) is shown as a function of the seven timegates chosen for inversion, (b) Three principal magnetic polarizabilities of a 81-mm mortar (ID:1339) is shown as a function of the seven timegates chosen for inversion. In case of both (a) and (b), we observe M_1 to be significantly larger than other two polarizabilities, and M_2 and M_3 are nearly identical. (c) Three principal magnetic polarizabilities of a clutter (ID:11) is shown as a function of the same seven timegates. 18

7 Feature Selection for TEMTADS using labeled training data 19

8 (a) The bar chart shows importance of the features in classifying the labeled training data for EM61-Array. The weights are normalized so that most important feature corresponds to unit weight. Note that only three features are "deemed" relevant for classification. (b) This 3x3 array of plots show pair-wise feature distribution for the EM61-array sensor, displaying TOI and clutter in red and blue, respectively. The diagonal plots show one-dimensional distribution along corresponding features. UXO and clutter are linearly separable if the blue and red histograms do not overlap. 20

9 (a) The bar chart shows importance of the features in classifying the training data for dual-mode MSEMS sensor. The plot shows twelve features (three magnetic polarizabilities for each of four timegates) extracted from the EM sensor, while the MAG sensor data has been rejected by SIG on grounds of being very noisy. (b) Pair-wise feature distribution for the MSEMS sensor is shown here for the two selected features. 21
10 (a) The bar chart shows importance of the features in classifying the MetalMapper training data. The plot shows fifteen features (three magnetic polarizabilities for each of five timegates) extracted from the sensor. (b) Pair-wise feature distribution for the MetalMapper sensor is shown here for the four selected features. 22
11 (a) This plot shows Receiver Operating Characteristics of RVM-based supervised classifier trained on labeled TEMTADS data. The percentage of UXO detected (Pd) is plotted as a function of the number of false alarms (Nfa) generated by the classifier. (b) This plot show the Risk Analysis, where the probability of detection (in red) and percentage excavation (in blue) are shown as a function of the relative cost of miss. The probability of detection achieves 100% at $C = 45$. SIG chose $C = 50$ for the testing data. (c) This plot shows the percentage of excavation for the testing data as a function of the threshold C. Note that SIG would propose to excavate 40% of the buried anomalies once it sets the threshold to 50. 24
12 Classification performance of SIG-extracted TEMTADS features 25
13 Classification performance of UBC-extracted TEMTADS features 26
14 Classification performance of the MetalMapper features 26
15 Classification performance of LBL-extracted BUD features 27
16 Classification performance of the MSEMS system using SIG-extracted features . . 28
17 Classification performance of the MSEMS system using UBC-extracted features . 28
18 Classification performance of the EM61-array using SIG-extracted features 29
19 Classification performance of the EM61-array using UBC-extracted features 30
20 Classification performance of the EM61-cart using UBC-extracted features 30
21 Classification performance of the SIG-extracted MAG features 31
22 Classification performance of the UBC-extracted MAG features 31
23 Performance of the actively learned classifier on TEMTADS and EM61-array data . 32
24 Itemized Cost for the SLO Data Analysis 33

Abbreviations Used

Abbreviation	Definition
11x	Buried object depth being 11 times its diameter
BUD	Berkeley UXO Discriminator
DoD	Department of Defense
DGM	Digital Geophysical Mapping
DSB	Defense Science Board
EMI	Electromagnetic Induction
ESTCP	Environmental Security Technology Certification Program
FAR	False Alarm Rate
FUDS	Formerly-Used Defense Site
GPS	Global Positioning System
IDA	Institute for Defense Analyses
LBL	Lawrence Berkeley Laboratory
MM	Munitions Management
MRS	Munitions Response Site
Pd	Probability of Detection
Pfa	Probability of false alarm
PNBC	Parametererized Neighborhood Based Classification
POC	Point of Contact
QA	Quality Assurance
QC	Quality Control
ROC	Receiver Operating Characteristic
RVM	Relevance Vector Machine
SERDP	Strategic Environmental Research and Development Program
SLO	San Luis Obispo
TEM	Time-domain Electro-Magnetic
TOI	Target of Interest
UBC	University of British Columbia
UXO	Unexploded Ordnance

Acknowledgments

SIG would like to acknowledge Dr. Erika Gasperikova of Lawrence Berkeley National Laboratory (LBL) and Dr. D. 'skip' Snyder for sharing the anomaly features extracted from the BUD and MetalMapper sensor, respectively. SIG would also like to acknowledge Prof. Doug Oldenburg of University of British Columbia and Dr. Steve Billings of Sky Research for sharing anomaly features extracted from TEMTADS, EM61-cart, EM61-array, Magnetometer, and MSEMS.

Executive Summary

SIG analysis of the anomaly data collected during the UXO discrimination study at Camp San Luis Obispo (SLO) performed well, especially for those anomalies for which the underlying dipole-model features were reliably extracted. Among the sensors deployed during this study, multi-axes and multi-coil electromagnetic sensors, such as TEMTADS, BUD, and MetalMapper, have produced superior classification performance in comparison to single-coil electromagnetic (EM61-cart, EM61-array) and magnetometer sensors. As proposed in the SLO demonstration plan [1], SIG has performed in-house feature extraction from the raw data collected by TEMTADS, Magnetometer, EM61-array and MSEMS dual-mode sensor, whereas features from the MetalMapper and the BUD sensor were provided by Skip Snyder of Snyder Geoscience and Erika Gasperikova of LBL, respectively. In addition, UBC/ Sky research provided us features for TEMTADS, Magnetometer, EM61-cart, EM61-array and MSEMS sensor, obtained using their own feature extraction model. This allowed SIG to compare the classification performance of two feature extraction schemes (SIG and UBC) under a single classifier framework. As mentioned in SIG memo [2] submitted to ESTCP program office in October'09, SIG has developed an improved forward model and associated feature extraction algorithm that enabled us to obtain fast and reliable estimate of the anomaly features. The improved model takes into account the exact shape of the transceiver coil while computing the voltage induced at the receiver. SIG trained a supervised and a semi-supervised classifier for each sensor-extractor pair, and provided IDA with a prioritized dig list corresponding to each classifier. The anomalies within a dig list were arranged in decreasing order of their likelihood of being a clutter. The items that could not be analyzed (*e.g.,* dipole model could not fit the data with acceptable accuracy) were appended at the end of the corresponding list. In addition, SIG has performed active learning on EM61-array and TEMTADS data. During active learning, SIG started with no *a priori* labeled data, and acquired labels for a small subset that are deemed "informative" by the active learning algorithm. SIG has provided two additional dig lists to IDA, corresponding to actively-learned classifiers for EM61-array and TEMTADS. The active learning algorithms produced superior classification performance using less labeled data than the passive (supervised and semi-supervised) approaches. The passive classifiers used a pre-assigned set of labeled data that were supplied to all data analysis groups for classifier training. SIG did not perform concept drift between the data collected at camp SLO and camp Sibert for two main reasons. The geophysical conditions between the two sites, Sibert and SLO, were quite different to pool the data together, and the new-generation sensors that produced high degree of discrimination between UXO and clutter during the SLO demonstration were not present during the Sibert data collection.

1 INTRODUCTION

1.1 BACKGROUND

The detection and remediation of unexploded ordnance (UXO) remains a high priority for the Department of Defense (DoD). According to 2003 Defense Science Board (DSB) report [3], the UXO cleanup problem is a very large-scale undertaking involving multi-million acres of land in more than fourteen hundred test sites scattered over the continental United States. It is estimated that only 10% of this area is occupied by UXO. The report estimates the cost of cleanup in the order of tens of billions in dollars. This cost consists of detection and classification of surface, buried, and partially buried UXO, followed by expensive excavation of these anomalies. The problem is aggravated by the fact that most of the UXO detection systems produce a large number of false alarms. This results in digging many holes where no UXO is present.

Significant progress has been made in the development of UXO detection and classification technologies over the past several years. However, most of the testing and validation of this technology (both in hardware designs and associated software algorithms) have been primarily limited to test sites with only limited application at live sites. Acceptance of these classification technologies required demonstration of system capabilities at real UXO sites under real world conditions. The FY06 Defense Appropriation contained funding for the "Development of Advanced, Sophisticated, Discrimination Technologies for UXO Cleanup" under the leadership of ESTCP, which responded by conducting a UXO Classification Study at the former Camp Sibert, AL [4]. The results of this demonstration were very encouraging. All of the classification approaches 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. However, the full capabilities of the developed systems could not be demonstrated due to limited availability of anomalies (a single target type, 4.2-inch mortar, was present in the demonstration site), and benign geology and topography of the site. To build upon the success of the first phase of this study, ESTCP sponsored a second study at the former Camp San Luis Obispo, CA, with more challenging topography and a wider mix of targets-of-interest (*e.g.,* 60mm, 81mm, 4.2-inch mortars, and 2.36-inch rockets).

1.2 OBJECTIVE OF THE DEMONSTRATION

As described in the ESTCP Unexploded Ordnance (UXO) Classification Study Plan [5], there were two primary objectives of the SLO demonstration. One objective was to test and validate UXO detection and discrimination capabilities of currently available and emerging technologies on a real site under operational conditions, and the other was to investigate the best possible methodology to implement the state-of-the-art technologies in cleanup operations involving live test sites.

The results of the SLO demonstration were very encouraging in the sense that a number of sensor systems deployed were capable of detecting all the buried UXO correctly with low false alarm rate. Signal Innovations Group (SIG) Inc. was involved in analyzing the UXO data collected at SLO and developing state-of-the-art classification algorithms that are optimized for the sensing framework. In accordance with the SIG demonstration plan [1], SIG has performed feature extraction using internally developed model-based inversion techniques for four sensors (EM61-

array, MAG, MSEMS and TEMTADS), whereas data collection teams provided features for other sensors (BUD, MetalMapper). This report will provide detailed description of the fast and robust feature extraction technique developed at SIG that works for both static cued systems (TEMTADS) and dynamic systems (EM61 and MAG). For a number of sensors (such as TEMTADS, MSEMS and EM61-array), SIG had two sets of features, one generated internally and the other provided by UBC/Sky research.

SIG has analyzed classification performance of each sensor-extractor pair using a supervised and semi-supervised classifier. Technical details of the two classifiers have been discussed in SIG Final Report on Sibert data analysis [4] (submitted in March 2008 and officially accepted by ESTCP), and they have also been published in peer-reviewed journals [6, 7]. Therefore, we shall discuss the implementational details briefly in this report, highlighting the salient features. A supervised classifier uses only a labeled feature set to train a classifier, while the semi-supervised approach utilized both labeled and unlabeled data. The performance of the multi-axis and/or multi-loop sensor systems (*e.g.,* TEMTADS, BUD and MetalMapper) were significantly better than their single-loop counterparts (*e.g.,* EM61-cart or EM61-array). SIG also observed superior classification performance of the time-domain electromagnetic systems (TEM) over the magnetometer sensor. In addition to developing static classification algorithms, SIG has also developed active learning classifiers for two of the sensors, TEMTADS and EM61-array. While actively-learned TEMTADS classifier outperformed the static supervised and unsupervised classifiers, the active EM61 classifier obtained performance similar to its supervised counterpart. However, it is worth noting that both of the actively learned algorithms used less labeled data than the non-active algorithms.

This report is organized in the following way. The technical aspects of our feature extraction and classification algorithms are explained briefly in section 2. The performance objectives are discussed in section 3. The site description and test design aspects are briefly mentioned in section 4 and 5, respectively. SIG was involved only in UXO discrimination analysis and had no role in data collection. Data analysis, including preprocessing, model parameter extraction and classifier training, is discussed in section 6. The analysis of the performance of various classifiers (supervised, semi-supervised and active learning) are presented in section 7. The cost assessments and implementation issues are discussed in section 8 and 9, respectively.

1.3 REGULATORY DRIVERS

SIG was involved in developing model-based techniques to discriminate buried UXO from clutter based on the responses collected by the sensor systems above the ground. The software development was off-line and did not come under any federal, state, or local regulations, assuming the data were collected satisfying all the existing laws of the land. Detailed description on the regulatory drivers and stakeholder issues can be found in ESTCP Demonstration Plan [5].

2 TECHNOLOGY - UXO DISCRIMINATION

Most of the current UXO detection systems can successfully detect buried UXO within 11x limit. However, these sensors also detect a lot of harmless cultural debris, or clutter, thereby increasing the overall cost of cleanup. It has been proven that many of these false alarms could be avoided if a model-based approach is employed to extract features from the raw data collected by these sensors. The physics-based models are designed to simulate typical UXO response, while clutter response does not fit well within the model.

2.1 FORMULATION OF THE FORWARD MODEL

SIG has been using a dipole model to approximate a buried UXO response [8] for last several years. This is a parametric model dependent on dipole location, orientation, and magnetic polarizability at all timegates. Reliable extraction of these features depend crucially on the size and quality of the available data, and the size of the parameter vector to be estimated. The parametric forward model used in previous studies was non-linear, and the inversion required a large number of random initializations to generate reliable estimate of the geophysical parameters of the buried anomalies. SIG has been using a limited number of timegates in order to obtain fast convergence of its inversion algorithm. In addition, sensors deployed during earlier demonstrations studies (Sibert, Fort Ord etc) collected response for a small number of timegates (*e.g.,* EM61 deployed in Sibert study used only four timegates).

Current research indicates that anomaly responses corresponding to late timegates provide important information that can distinguish UXO from clutter. This led to the development of next-generation sensor systems, such as TEMTADS and MetalMapper, that obtain anomaly response at hundreds of timegates. Faced with the increased capability of these sensors, SIG has significantly improved the forward model and the associated inversion algorithm. The improvements are as follows.

- The magnetic field produced by the transmitting coil is currently described in a closed form (using Biot-Savart law). Therefore, SIG can incorporate the exact shape of the transmitting coil in its forward model, rather than approximating it as a dipole (as done during the Sibert analysis). Similarly, the voltage induced in the receiving coil is determined via the principle of reciprocity, and no dipole approximation of the receiver loop is introduced as well.

- The orientation angles of a buried anomaly (nonlinear parameters of the dipole model) are not extracted directly from the measured field data. Instead, we integrate them with the magnetic polarization dyadic and extract six independent elements for each timegate (for a symmetric 3x3 polarization matrix). Thus, the number of nonlinear parameters of the dipole model is reduced, thereby increasing the speed of convergence by reducing the generation of local solutions.

- Three principal magnetic polarizabilities and three orientation angles can be readily evaluated in a post-processing step after the feature inversion algorithm has converged. The sep-

aration of feature inversion and principal polarizability evaluation step enabled us to obtain quick and robust estimation of the anomaly parameters.

It is well known in the UXO community that electromagnetic induction (EMI) is one of the most promising approaches to discriminate UXO from metallic clutters. A standard EMI system is generally consisting of one or multiple transmitting and receiving coils. In a sensor of this type, a buried anomaly (UXO or metallic clutter) is excited by the primary magnetic field produced due to a series of current pulses flowing through the transmitting coils. The resulting eddy currents produced within the anomaly generates a secondary magnetic field that induces a voltage in the receiving coils. If the anomaly dimensions are relatively small compared to its distance from the sensing coils, the secondary magnetic field can be approximated as being generated by a point source (dipole) located at the center of the anomaly. The strength, or dipole moment of the point source is a function of the geophysical parameters (shape, size, metallic content, and orientation) of the buried anomaly.

Assume that three principal axes of the induced magnetic dipole moment are co-located with a specific coordinate system, and the magnetic polarization dyadic (for each timegate) can be written as follows.

$$\bar{\bar{\boldsymbol{P}}} = \begin{bmatrix} P_{x'} & 0 & 0 \\ 0 & P_{y'} & 0 \\ 0 & 0 & P_{z'} \end{bmatrix} \tag{1}$$

Let us assume that the magnetic field at the center of the anomaly due to the current in the transmitting coil can be described as $\vec{H}_t = [H_x\, H_y\, H_z]^T$ (presented in transmitting coil coordinate system), where the superscript T and subscript t represent "transpose" and "transmitting", respectively. Generally, the coil coordinate system is not co-located with the anomaly coordinate system. Therefore, the primary magnetic field in the object coordinate system can be obtained through a rotation transform as

$$\vec{H}_p = R\vec{H}_t, \tag{2}$$

where R is a rotation matrix represented as

$$R = \begin{bmatrix} \cos\psi\cos\theta\cos\phi - \sin\psi\sin\phi & \cos\psi\cos\theta sin\phi + \sin\psi\sin\phi & -\cos\psi\sin\theta \\ \cos\psi\cos\theta\cos\phi - \sin\psi\sin\phi & \cos\psi\cos\theta sin\phi + \sin\psi\sin\phi & -\cos\psi\sin\theta \\ \sin\theta\cos\phi & \sin\theta\sin\phi & \cos\theta \end{bmatrix}, \tag{3}$$

where θ represents the angle between the vertical (z) axes of the two coordinate systems, and ϕ represents the angle between the projection of vertical axis of the anomaly coordinate system onto the horizontal plane of the transmitting coil coordinate system and x-axis of the coil coordinate system. The angle ψ represents the rotation angle about the z-axis of the object coordinate system. The magnetic dipole moment $\vec{M}$, induced at the object center, is proportional to the product of the magnetic polarization dyadic and the primary magnetic field.

$$\vec{M} = \bar{\bar{P}}\vec{H}_p = \bar{\bar{P}}R\vec{H}_i \tag{4}$$

Assuming that the receiving coil coordinate system is co-axial with the transmitting coil system, the magnetic dipole moment $\vec{m}$ induced at the receiving coil coordinate system can be written as

$$\vec{m} \propto R^T \bar{\bar{P}} R \vec{H}_i, \tag{5}$$

and the magnetic field created at the center of the receiver coil (at a distance r from the anomaly dipole) can be evaluated in terms of the dipole moment as

$$\vec{H}_r = \frac{1}{4\pi} \frac{3\vec{r}(\vec{m} \cdot \vec{r}) - \vec{m} r^2}{r^5}. \tag{6}$$

In previous analyses (such as Sibert) SIG has modeled both the transmitting and receiving coils as magnetic dipoles located at the center of the loop. In such a case, the above expression evaluates the magnetic field at the center of the receiver coil, and hence the voltage (proportional to the magnetic field) induced at the receiver.

For the San Luis Obispo data analysis, SIG has developed a more accurate model. According to the principle of reciprocity, the voltage V_r, induced in a receiving coil is proportional to

$$V_r \propto \vec{H}_r \cdot \vec{m}, \tag{7}$$

where $\vec{H}_r$ is the magnetic field that would be produced at the center of the anomaly (location of the dipole) by an unit current flowing in the receiving coil. In order to find the voltage induced in the receiving coil, one needs to find the magnetic field intensity $\vec{H}_r$ at the object center, and the magnetic dipole moment $\vec{m}$ of the anomaly induced due to the primary magnetic field. The evaluation of magnetic dipole $\vec{m}$ due to the incident magnetic field $\vec{H}_i$ has been described in Eq. 5. The following section describes the evaluation of magnetic field at a point due to an unit current in a rectangular loop (utilized for calculating both $\vec{H}_i$ and $\vec{H}_r$).

2.2 ANALYTICAL COMPUTATION OF THE MAGNETIC FIELD

Assume that a rectangular coil is placed in the x-y plane and centered at the origin, with side dimensions 2a and 2b, as shown in Figure 1. The vector potential due to current I flowing counterclockwise through the coil can be written as

$$\vec{A} = A_x \hat{x} + A_y \hat{y} = \frac{\mu_0 I}{4\pi} \left[\left(\int_{l=-a}^{l=a} \frac{dl}{R_1} + \int_{l=a}^{l=-a} \frac{dl}{R_2} \right) \hat{x} + \left(\int_{l=-b}^{l=b} \frac{dl}{R_3} + \int_{l=b}^{l=-b} \frac{dl}{R_4} \right) \hat{y} \right] \tag{8}$$

where

$$\begin{aligned} R_1 &= \sqrt{(x-l)^2 + (y+b)^2 + z^2} \\ R_2 &= \sqrt{(x-l)^2 + (y-b)^2 + z^2} \\ R_3 &= \sqrt{(x-a)^2 + (y-l)^2 + z^2} \\ R_4 &= \sqrt{(x+a)^2 + (y-l)^2 + z^2} \end{aligned}$$

The above integrations can be evaluated analytically. For example, the x component of the vector potential can be written as

$$A_x = \frac{\mu_0 I}{4\pi} \ln \left[\frac{(r_1 + a + x)}{(r_2 - a + x)} \cdot \frac{(r_3 - a + x)}{(r_4 + a + x)} \right], \tag{9}$$

where $r_i, i = 1, 2, 3, 4$ are the distances of four corners of the coil from the observation point P. The y component of the vector potential can be evaluated in a similar fashion.

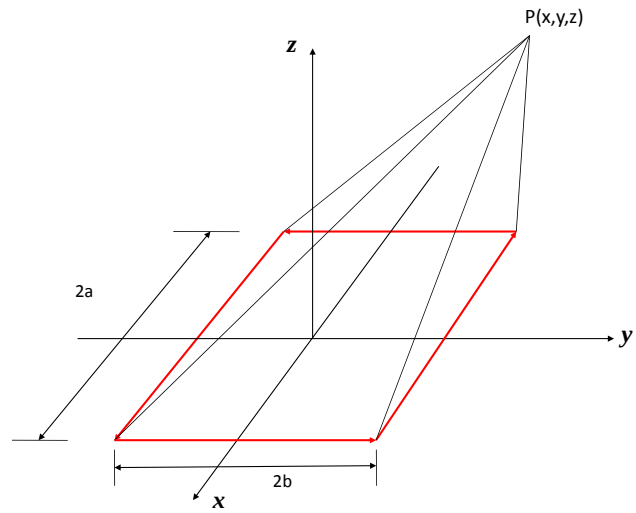


Figure 1: The figure shows the geometry of a rectangular coil (shown in red) in the x-y plane. Our objective is to obtain an analytical expression of the magnetic field at any location P(x,y,z) due to an unit current flowing through the rectangular coil.

The magnetic field produced by the coil can be readily obtained according to Maxwell's equations as

$$\vec{H} = \frac{1}{\mu_0} \nabla \times \vec{A},$$

and the three components of the magnetic field can be evaluated in closed form as shown below.

$$\begin{aligned}
H_x &= \frac{1}{\mu_0} \frac{\partial A_y}{\partial z} = \frac{I}{4\pi} \sum_{i=1}^{4} \frac{(-1)^{i+1} z}{r_i (r_i + d_i)} \\
H_y &= \frac{1}{\mu_0} \frac{\partial A_x}{\partial z} = \frac{I}{4\pi} \sum_{i=1}^{4} \frac{(-1)^{i+1} z}{r_i \left[r_i + (-1)^{i+1} c_i \right]} \\
H_z &= \frac{1}{\mu_0} \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right) = \frac{I}{4\pi} \sum_{i=1}^{4} \left[\frac{(-1)^i d_i}{r_i \left[r_i + (-1)^{i+1} c_i \right]} - \frac{c_i}{r_i (r_i + d_i)} \right]
\end{aligned}$$

where

$$\begin{aligned} c_1 &= -c_4 = a + x \\ c_2 &= -c_3 = a - x \\ d_1 &= d_2 = y + b \\ d_3 &= d_4 = y - b \end{aligned}$$

The magnetic fields $\vec{H}_r$ and $\vec{H}_i$ mentioned in the previous subsection can be calculated in closed forms using the above equations. Therefore, dipole approximations of the transmitting and receiving coils are removed, since the current formulation can handle the exact shape of these coils while calculating the voltage induced in the receiving coils.

2.3 SUPERVISED BAYESIAN CLASSIFICATION

The Bayesian supervised classifier [6] employs the functional

$$c(\boldsymbol{f}) = \sum_{n=1}^{N} w_n K(\boldsymbol{f}, \boldsymbol{b}_n) + w_0 = \boldsymbol{w}^{\mathrm{T}} \phi(\boldsymbol{f}) \tag{10}$$

where $K(\boldsymbol{f}, \boldsymbol{b}_n)$ is termed a "kernel" [6], quantifying the similarity between feature vector $\boldsymbol{f}$ and the nth basis vector $\boldsymbol{b}_n$, $\boldsymbol{w} = \{w_0, w_1, w_2, \ldots, w_N\}^{\mathrm{T}}$, and $\phi(\boldsymbol{f}) = \{1, K(\boldsymbol{f}, \boldsymbol{b}_1), K(\boldsymbol{f}, \boldsymbol{b}_2), \ldots, K(\boldsymbol{f}, \boldsymbol{b}_N)\}^{\mathrm{T}}$. The basis vectors $\boldsymbol{B} = \{\boldsymbol{b}_n\}_{n=1,N}$ are typically defined in terms of the available training data, using clustering techniques. The RVM employs a Bayesian approach, where the probability of feature vector $\boldsymbol{f}$ to be associated with label $l \in \{-1, 1\}$, corresponding to UXO/non-UXO, is defined by the logistic link function

$$p\left(l = +1 | \boldsymbol{f}, \boldsymbol{D}_L, \boldsymbol{w}\right) = \sigma\left[c(\boldsymbol{f})\right] = \frac{1}{1 + \exp\left[-c(\boldsymbol{f})\right]} \tag{11}$$

where $\boldsymbol{D}_L$ represents the (labeled) training data $\{\boldsymbol{f}_n, l_n\}_{n=1,N_L}$.

The classifier treats the weights as random variables, with statistics dependent on the properties of the labeled training data $\boldsymbol{D}_L$. Within a Bayesian framework, our approach tries to estimate the full posterior density of the classifier weights, given the labeled training data $\boldsymbol{D}_L$, rather than estimating a single *most likely* estimate of the weight vector. In particular, the weights are characterized by $p(\boldsymbol{w}|\boldsymbol{D}_L, \boldsymbol{\alpha})$, where the $N + 1$ dimensional vector $\boldsymbol{\alpha}$ (one component of $\boldsymbol{\alpha}$ is associated with each component of $\boldsymbol{w}$) defines the parameters associated with the density function for $\boldsymbol{w}$. The classifier is "regularized" by introducing a prior density function for the parameters $\boldsymbol{\alpha}$, with which one may impose desirable constraints on the form of $p(\boldsymbol{w}|\boldsymbol{D}_L, \boldsymbol{\alpha})$. Recognizing that sparseness is a desirable property of the final classifier, the density function associated with $\boldsymbol{\alpha}$, denoted $p(\boldsymbol{\alpha})$, is designed to impose a sparse solution (most weights in $\boldsymbol{w}$ being zero). The final probability calculated for classification is represented as

$$p(l = +1|\boldsymbol{f}, \boldsymbol{D}_L) = \iint p(l = +1|\boldsymbol{f}, \boldsymbol{D}_L, \boldsymbol{w}) p(\boldsymbol{w}|\boldsymbol{D}_L, \boldsymbol{\alpha}) p(\boldsymbol{\alpha}) d\boldsymbol{w} d\boldsymbol{\alpha} \tag{12}$$

An approximate iterative procedure [6] has been developed for computation of the integrals in (12), yielding a relatively simple algorithm.

2.4 SEMI-SUPERVISED LEARNING

Semi-supervised learning is applicable to any sensing problem for which all of the unlabeled data are available at the same time, and therefore, this approach is applicable to most wide-area sensing problems, particularly for the current demonstration study. In most practical applications, semisupervised learning has been found to yield superior performance relative to the widely applied supervised algorithms. Suppose $\boldsymbol{G} = (\boldsymbol{\mathcal{F}}, \boldsymbol{\mathcal{W}})$ is a graph where $\boldsymbol{\mathcal{F}} = \{\boldsymbol{f}_1, \ldots \boldsymbol{f}_N\}$ represents a set of N nodes (these are our feature vectors, one for each anomaly, labeled and unlabeled), and $\boldsymbol{\mathcal{W}} = [w_{ij}]_{N\times N}$ is an affinity matrix with w_{ij} indicating the similarity between the feature vectors $\boldsymbol{f}_i$ and $\boldsymbol{f}_j$. Considering a radial basis function to define affinity, we have

$$w_{ij} = \exp\left(-\frac{||\boldsymbol{f}_i - \boldsymbol{f}_j||^2}{2\sigma^2}\right) \tag{13}$$

where $||\cdot||$ represents the Euclidean norm. Based on the above formulation, one can design a Markov transition matrix $\boldsymbol{A} = [a_{ij}]_{N\times N}$, where

$$a_{ij} = \frac{w_{ij}}{\sum_{k=1}^{N} w_{ik}} \tag{14}$$

represents the probability of transitioning from node $\boldsymbol{f}_i$ to $\boldsymbol{f}_j$. Any two nodes (*i.e.,* feature vectors) $\boldsymbol{f}_i$ and $\boldsymbol{f}_j$ are said to be neighbors, if $a_{ij} > 0$ and $\mathcal{N}(\boldsymbol{f}_i) = \{\boldsymbol{f} : \boldsymbol{f} \sim \boldsymbol{f}_i\} \subseteq \mathcal{F}$ represents the neighborhood of the feature vector $\boldsymbol{f}_i$. The semi-supervised binary classifier with label $y \in \{-1, 1\}$ uses the functional

$$p(y_i|\mathcal{N}(\boldsymbol{f}_i), \boldsymbol{\theta}) = \sum_{j=1}^{N} a_{ij} p(y_i|\boldsymbol{f}_j, \boldsymbol{\theta}) \tag{15}$$

Assuming $\mathcal{L} \subseteq \{1, 2, \ldots, N_L\}$ representing the set of labeled data indices, the likelihood functional can be written as

$$p(\{y_i, i \in \mathcal{L}|\mathcal{N}(\boldsymbol{f}_i), \boldsymbol{\theta}) = \prod_{i\in\mathcal{L}} p(y_i|\mathcal{N}(\boldsymbol{f}_i), \boldsymbol{\theta}) = \prod_{i\in\mathcal{L}} \sum_{j=1}^{N} a_{ij} p(y_i|\boldsymbol{f}_j, \boldsymbol{\theta}) \tag{16}$$

Estimation of classifier parameters $\boldsymbol{\theta}$ can be achieved by maximizing the log likelihood via an Expectation-Maximization algorithm [7]. To enforce sparseness of $\boldsymbol{\theta}$ (enforcing most of the components of the parameter vector $\boldsymbol{\theta}$ to be zero), one may impose a zero-mean Gaussian prior on each dimension of $\boldsymbol{\theta}$.

2.5 ACTIVE LEARNING

The main objective of the current demonstration is to test and validate the efficacy of state-of-the-art sensors and associated statistical analysis techniques in correctly detecting and excavating as high a percentage of UXOs as possible while leaving as much of the non-UXOs as possible unexcavated. Given that all the anomalies will be eventually excavated for this demonstration, one may ask in which order these anomalies should be excavated that improves the performance of the classifier algorithm most, and thereby enhancing our confidence on the estimated identity of the

anomalies that are yet to be excavated. Specifically, one may ask which unlabeled sensor signature would be most informative to improve classifier performance if the associated label could be made available. It has been shown [9] that this question can be answered in a quantitative information-theoretic manner.

For active label selection, posterior distribution of the classifier is approximated as a Gaussian distribution with

$$p(\boldsymbol{\theta}|\boldsymbol{D}_L) \approxeq \mathcal{N}(\boldsymbol{\theta}|\hat{\boldsymbol{\theta}}, \mathrm{H}^{-1}) \tag{17}$$

where $\hat{\boldsymbol{\theta}}$ is the *maximum a posteriori* (MAP) estimate of the classifier, as derived in the semi-supervised classifier design and H represents the posterior precision matrix $\mathrm{H} = \nabla^2(-\log p(\boldsymbol{\theta}|\{y_i, \mathcal{N}(\boldsymbol{f}_i) : i \in \mathcal{L})$. With one more data point $\boldsymbol{f}_{i*}$ with label y_{i*} as the next available labeled data, the posterior precision changes to

$$\mathrm{H} = \sum_{i' \in \mathcal{L} \cup \{i*\}} \sum_{j=1}^{N} \delta_{i'j} p(y_{i'}|\boldsymbol{f}_j, \boldsymbol{\theta}) \times (1 - p(y_{i'}|\boldsymbol{f}_j, \boldsymbol{\theta}))\boldsymbol{f}_j \boldsymbol{f}_j^{\mathrm{T}} - \nabla^2 \ln \gamma(\hat{\boldsymbol{\theta}}) \tag{18}$$

which can be further simplified by assuming $\delta_{ij} = 1$ if $i = j$, and 0 if $i \neq j$. The new precision matrix can be written as

$$\mathrm{H}' = \mathrm{H} + p(y_{i*}|\boldsymbol{f}_{i*}, \boldsymbol{\theta})(1 - p(y_{i*}|\boldsymbol{f}_{i*}, \boldsymbol{\theta}))\boldsymbol{f}_{i*}\boldsymbol{f}_{i*}^{\mathrm{T}} \tag{19}$$

Our objective is to choose the feature vector for labeling that maximizes the mutual information (I) between the classifier $\boldsymbol{\theta}$ and the new data point to be labeled. The mutual information can be quantified as the expected decrease of the entropy of $\boldsymbol{\theta}$ after $\boldsymbol{f}_{i*}$. and y_{i*} are observed.

$$I = \frac{1}{2} \log \frac{|\mathrm{H}'|}{|\mathrm{H}|} = \frac{1}{2} \log\{1 + p(y_{i*}|\boldsymbol{f}_{i*}, \boldsymbol{\theta}) \times [1 - p(y_{i*}|\boldsymbol{f}_{i*}, \boldsymbol{\theta})]\, \boldsymbol{f}_{i*}^{\mathrm{T}} \mathrm{H}^{-1} \boldsymbol{f}_{i*}\} \tag{20}$$

It is important to note that the mutual information I is large when $p(y_{i*}|\boldsymbol{x}_{i*}, \boldsymbol{\theta}) \approx 0.5$. Hence, the active learning prefers to acquire labels on those unlabeled samples for which the current classifier is most confused or uncertain.

2.6 ADVANTAGES AND LIMITATIONS OF TECHNOLOGY

An overarching problem in UXO remediation is that almost 90% of the remediation cost comes from excavating scrap metal and other clutter items. This false-alarm problem dramatically increases costs and slows remediation. Standard mag-and-flag techniques are particularly prone to the false alarm problem, since any ferrous subsurface anomaly is flagged for excavation. Modern techniques incorporate features associated with inversion of physics-based models and apply these features in a statistical decision framework. Current physics-based statistical approaches are limited by a lack of site-specific training data. It should also be noted that in many cases discrimination has not been attempted at all, or has been performed poorly, due to inappropriate data quality (*e.g.,* low SNR).

A significant benefit to the DoD of the proposed project is the reduced number of total excavations required to clean a given site. Specifically, most of the false alarms associated with current tech-

nologies are attributed to the mismatch between available labeled training data and the unlabeled data to be classified. Active learning constitutes a state-of-the-art mathematically rigorous means of adaptively augmenting the training set, based on the site-dependent observed data. The active-learning framework provides two dig lists: the first defines a set of signatures for which access to the associated labels is most informative to classifier design, and after these labels are acquired a second dig list is provided specifically targeted toward excavating UXO.

Another key component of the technology demonstrated and evaluated here is semi-supervised learning, in which classification of any given target is placed in the context of all targets of interest from a given site. As demonstrated in prior SERDP results, active and semi-supervised learning have resulted in significantly reduced total excavations, while achieving high UXO detection.

While the goal was to dramatically reduce the number of false alarms associated with a clean-up activity, it would be disingenuous to suggest that 100% of the UXO can be detected with no false alarms. There are certainly classes of clutter and geophysical anomalies that cannot be reliably differentiated from UXO, possibly as a result of data quality, or as a result of an inherent overlap of the two classes of objects in the feature space. While the techniques demonstrated on SLO data have reduced the false alarm problem, it did not mitigate the problem completely.

3 PERFORMANCE OBJECTIVES

The performance objectives of the demonstration are summarized in Figure 2. There were eleven sensor arrangements (each sensor-extractor combination is treated as one arrangement), each of which resulted in two different dig lists (one for a supervised classifier and one for a semi-supervised classifier). In addition, SIG has developed two additional dig lists for TEMTADS and EM61-array using active learning experiments. As mentioned above, SIG decided to not perform concept drift between Sibert and SLO data due to their geophysical differences and non-availability of next-generation sensor data from Sibert.

SIG proposed to perform model-based feature extraction for Magnetometer, EM61-array, TEMTADS and MSEMS sensor, whereas BUD and MetalMapper features were provided to us by LBNL and Snyder Geoscience, respectively. SIG proposed to extract reliable features for each of the sensors which are consistent with the ground truth. SIG also proposed to monitor the "goodness of fit" of the simulated response and only subjected the "good" features to subsequent classification. Our objective was to obtain reliable target parameters for more than 90% of the anomaly responses.

Each sensor-extractor pair produced two dig lists, one for the supervised and one for the semi-supervised classifier. Risk analysis was used to determine the division between items to declare UXO and items to declare non-UXO. Performance of each sensor-extractor pair was based on the results of this dig-list declaration. Sensor performance comparison (in terms of probability of detection and number of false alarms) was based on these results.

Furthermore, the classifiers that exploit active learning was compared to the corresponding classifiers that employ traditional learning formulations (*i.e.,* with a predefined training set). It was

Performance Objective	Performance Metric	Data Required	Expected Results
Correct characterization of dipole model parameters	Accuracy of estimated target parameters	*M* (dipole moment strength) ω(dipole decay constants) x, y, z (position)	Relative values of parameters comport with size and identity, and the target position from truth
Maximize correct classification of munitions	Number of targets-of-interest retained	Prioritized anomaly lists for each sensor data set	Discrimination retains all detected objects of interest above the no-dig threshold
Maximize correct classification of non-munitions	Number of false alarms eliminated	Prioritized anomaly lists for each sensor data set	Reduction of false alarms by >30% while retaining all targets of interest
Demonstrate appropriate threshold can be specified via risk analysis	Pd and False Alarm rate at the specified operating threshold	Specify two threshold to identify "high confidence" munitions and "high confidence" non-munitions	>30% of non-munitions can be correctly labeled
Minimize number of anomalies that cannot be analyzed	Number of anomalies that cannot be analyzed	Target parameters	Reliable target parameters can be estimated for >90% of anomalies on each sensor's detection list.

Figure 2: Performance objectives for the Camp San Luis Obispo study

expected that performance with active learning will be superior to the non-active-learning (in the sense of requiring less labeled data).

SIG has examined the validity of the dipole model inversion process by comparing the fit model parameters against the size and identity of the items. For example, the dipole strength of a large item should have a larger value. Direct comparisons between the fit model parameters and the true parameters were not possible because the latter are not available. In order to obtain a more general measure of performance, ROC curve analysis was utilized.

4 SITE DESCRIPTION

4.1 TEST SITE SELECTION

The ESTCP UXO Classification Study Demonstration site has been selected to be former Camp San Luis Obispo, California. The 11.8 acre demonstration area was chosen based on the results of the initial EMI survey [10]. The site was a mortar target range situated on a hilltop in Munitions Response Site (MRS) [10], approximately five miles northwest of San Luis Obispo, CA. This demonstration area 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 (4.2-inch mortar), and item "size" was an effective discriminant. At this site, there are at least four targets-of-interest: 60-mm, 81-mm, and 4.2-in mortars and 2.36-in rockets. This introduces another layer of complexity into the process.

4.2 TEST SITE HISTORY AND CHARACTERISTICS

Camp San Luis Obispo was established in 1928 by the state of 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 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. 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 HRR identified eighteen ranges and sixteen training areas present at Camp San Luis Obispo during the Korean War. Following the Korean War, the camp was maintained in inactive status until it was relinquished by the Army in the 1960s and 1970s. 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.

The Camp San Luis Obispo site consists mainly of mountains and canyons classified as grassland, wooded grassland, woodland, or brush. Los Padres National Forest (woodland) is located to the north-northeastern portion of the site. 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. 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.

The demonstration site is configured as one 11.8-acre area. It spans the hillside that is the historical mortar target. The cart systems surveyed 45 30×30 square-meter grids within this area for a total of 10 acres. The vehicular systems surveyed the entire area. The calibration strip and training pit were located off the site, convenient to the access road. Details of the final site is outlined in the ESTCP UXO classification Study Plan.

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 survey:

- 4.2-inch white phosphorus mortar
- 4.2-inch base plate
- 3.5-inch rocket
- 37mm mortar
- 75mm
- 105mm
- 60mm mortar
- 81mm mortar
- practice bomb and
- 30 cal casings and fuzes.

At the particular site of this demonstration, 60-mm, 81-mm, 4.2-in mortars and mortar fragments, and 2.36-in rockets have been observed. The excavation of two grids as part of the preparatory activities has confirmed these observations and provided information on the depths of munitions at this target site.

5 TEST DESIGN

SIG was involved in data discrimination phase of the demonstration study and had no role to play in data collection. Information regarding the site preparation, system specification, calibration activities, data collection and validation should be accessed from the ESTCP demonstration plan [5] and final reports submitted by various data collection teams.

6 DATA ANALYSIS

6.1 PROCEDURE FOR DATA CHIP SELECTION

Section 2 has described an accurate forward model that SIG has developed for explaining the responses collected by the above-ground sensor systems. This section will describe the procedure SIG followed to identify good quality data that is suitable for model-parameter inversion. As mentioned before, SIG has performed feature extraction for both dynamic (MAG, EM61-array and MSEMS) and cued (TEMTADS) sensors. In case of a dynamic sensor in survey mode, the sensor is moved in a linear track while the receiver is collecting data at regular intervals. Once the entire area has been surveyed and geophysical mapping has been performed, specific sensing locations are “deemed” anomaly detections using a “mag and flag” technique. For each of the survey-mode

sensors, SIG was provided with geo-located response over a large area, along with a master list of anomaly locations detected by that sensor.

For each anomaly on the master list, SIG identified the anomaly location in GPS coordinate and extracted a 5m x 5m data chip centered around that location. We first ascertained whether the chosen square area includes any other detected anomalies, according to the list of anomaly locations provided. The received response at the first time gate were inspected visually to detect the overall quality of the received response. First timegate was chosen for visual inspection since it corresponds to highest signal strength. In case of MAG, we displayed the magnitude image over the square area.

Visual inspection identifies the number of anomalies close to the "detected" location (according to the master list). Assuming the image boundary of the associated anomaly is clear, we draw a polygon along the boundary. If there are multiple anomalies within the polygon, then we extract features for all the anomalies jointly and subsequently classify each of the anomalies as UXO or clutter. However, we assign a single rank to the entire set based on the likelihood of the most-likely UXO among all members. SIG has developed a matlab-based interactive algorithm to draw polygon around a detected anomaly by specifying the node locations of the polygon and the underlying code extracted the survey locations that were contained within the polygon. Data corresponding to these locations were used for subsequent feature extraction. Figure 3 shows cases where there are are one, two or three anomalies within the manually selected polygon (shown in red). The underlying Matlab code extracts all survey data contained within the polygon to fit one or multiple buried anomalies (as identified during visual inspection).

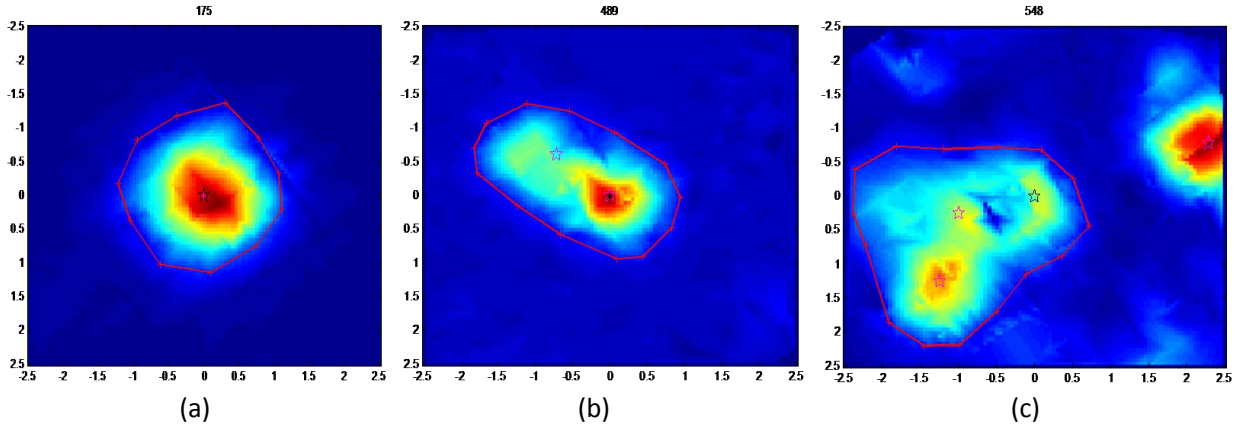


Figure 3: (a) The figure shows a 5mx5m data chip centered around anomaly ID 175. The drawn polygon is visually inspected to contain only one anomaly, (b) The figure shows 5mx5m data chip centered around anomaly ID 489. The drawn polygon contains two anomalies. The data contained within the polygon is subjected to feature extraction model to obtain two sets of anomaly parameters, (c) The figure shows 5mx5m data chip centered around anomaly ID 548. Three sets of anomaly parameters are obtained jointly from the data contained within the polygon.

There were a few cases where SIG could not clearly isolate the data that corresponds to the anomaly location given in the master list. Figure 4 shows three distinct cases where SIG could not extract reliable data for feature extraction. In Figure 4(a), the detected location (shown as green circle) is close to very strong signal making it hard to identify the associated data points. There are too many targets (> 3) within the anomaly boundary in Fig. 4(b), and low SNR in Figure 4(c). These cases were decided by SIG as "Can't analyze" cases and appended at the bottom of the prioritized dig lists delivered to IDA.

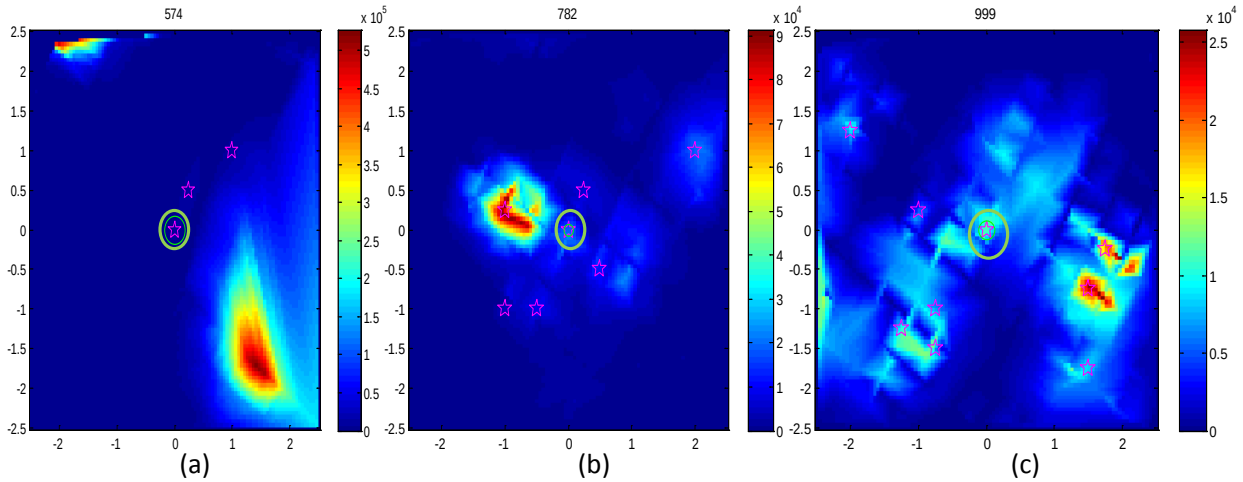


Figure 4: (a) The figure shows 5mx5m data chip centered around anomaly ID 574. The anomaly is too close to very strong signal from a neighboring anomaly, (b) The figure shows 5mx5m data chip centered around anomaly ID 782. There are too many anomalies in close proximity, making it difficult to extract their parameters reliably, (c) The figure shows 5mx5m data chip centered around anomaly ID 999. Signal-to-noise ratio is considered too low for reliable feature extraction.

SIG has also performed data selection for a cued sensor, TEMTADS. In the cued mode, a sensor is supplied with a list of detected anomaly locations (using a survey sensor like EM61 or MAG). The TEMTADS system was driven to each of these locations and placed on top of the detection location with the help of a GPS-guided system. TEMTADS consists of 25 transmitters and receivers arranged in a 5x5 horizontal array spanning a 2mx2m square area. Once the system is placed exactly above the buried anomaly, each of the 25 transmitters are excited in sequence and responses are collected at all 25 receivers. For each anomaly on the master list, SIG was provided with response data for all 625 Tx-Rx combinations collected at 115 timegates. Given a large horizontal span of the TEMTADS array, SIG has concluded that data set could be pruned further to identify only those combinations that contain high quality data. Rather than visual inspection of all 625 channels, we displayed only monostatic responses received at 25 (5x5) channels. We visually selected a subset of these 25 Tx/Rx locations that correspond to high to moderate signal strength. A subset of chosen locations (say 15) corresponds to 225 combinations of Tx-Rx pairs that could be used for feature extraction. SIG had decided to further prune the dataset by choosing the top 100 channels based on the first timegate signal strength. Note that the last pruning step is automated whereas the visual inspection based channel selection was done manually. SIG feels that manual

selection was crucial to identify the location of the anomaly on the received response map (as they were not always co-located with the master list locations), and carefully ignore channel that might have captured signal from an adjacent anomaly.

6.2 MODEL PARAMETER EXTRACTION

Based on the generalized forward model described above, the three principal polarization terms in the magnetic polarization dyadic, the orientation angles in the rotation matrix, as well as the position of the object can be extracted directly from the field data by using nonlinear least-square solver. However, it is well known that the trigonometric functions in the rotation matrix are non-linear multi-valued functions. Therefore, direct extraction of the orientation angles might be easily trapped by one of the many possible local solutions.

On the other hand, consider the general magnetic polarization dyadic tensor which is defined as

$$\bar{\bar{M}}_g = R^T \bar{\bar{M}} R$$

One can easily verify that the tensor is a symmetric matrix, and therefore, there are only six independent elements (parameters). All the diagonal terms of the general magnetic polarization dyadic are non-negative, and the general magnetic polarization dyadic is a linear function of the measured field data. We can exploit this by extracting the six parameters directly from the measured data (using least-square inversion), rather than extracting the three principal polarization components and three rotation angles directly. the number of non-linear parameters in the model are greatly reduced in this implementation, and the problem of local solutions are significantly relieved.

Once six parameters of the general dyadic tensor are extracted, the three orientation angles as well as the three principal polarization components can be readily determined. It can be easily derived that three principal polarization components correspond to the three eigenvalues of the general magnetization dyadic, and the associated unitary matrix is equivalent to the rotation matrix R. Therefore, the three orientation angles can be extracted directly from the rotation matrix as follows.

$$\begin{aligned}
\theta &= \cos^{-1} R_{33} \\
\psi &= \sin^{-1}\left(\frac{R_{23}}{\sin\theta}\right) \\
\phi &= \sin^{-1}\left(\frac{R_{12}}{\cos\psi\cos\theta + \sin\psi}\right)
\end{aligned}$$

In summary, the response of the buried object can be modeled by a set of parameters: three position parameters associated with the object location, and six parameters associated with the general magnetic polarization tensor for each timegate. The principal magnetic polarization components and the object's orientation angles can be further determined through post-processing these parameters. It should be noted that SIG did not assume the buried objects to be symmetrical for the SLO data analysis.

Figure 5 shows a typical example of the convergence of the new approach. The figure on the left shows the three principal polarizabilities (in blue, red and green) of a 81mm mortar (ID:1081),

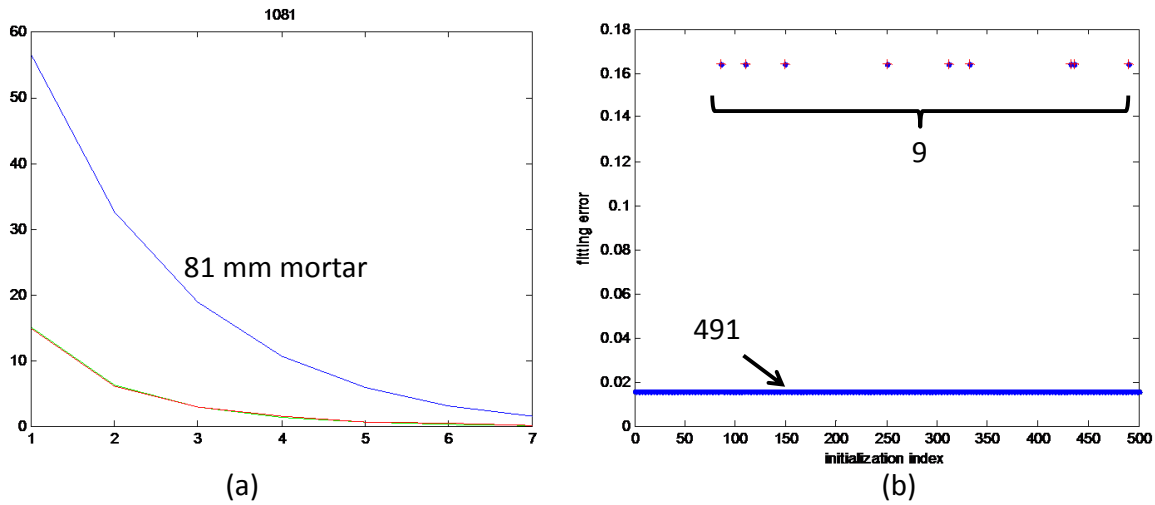


Figure 5: (a) Three principal magnetic polarizabilities of a 81mm mortar (ID:1081) is shown as a function of the seven timegates chosen for inversion. We observe M_1 to be significantly larger than other two polarizabilities. M_2 and M_3 are nearly identical. The response is expected typically of a cylindrical anomaly (81mm mortar is a cylindrical UXO), (b) The plot shows the fitting error of 500 random initializations for anomaly ID 1081 (a 81mm mortar). We observe good fit with the measured field data for 491 initiations, and all those 491 initializations converged to the parameter vector shown in (a).

as a function of the timegates used for feature extraction. Note that the object's cylindrical shape can be easily verified by one predominantly large and the other two nearly identical magnetic polarizabilities. The figure on the right shows the plot of fitting error for a set of 500 random initializations. Note that 491 initializations resulted in good fit (low fitting error). All good fits correspond to the same converged parameter vector. We have observed similar performance for almost all anomalies having moderate/low background noise in absence of strong response from nearby anomalies.

Figure 6 shows the plot of the converged magnetic polarizability parameters for two types of UXO (4.2-inch mortar in (a) and 81-mm mortar in (b)), and a rock (clutter). The picture of the UXO are shown inset. Note that both of the UXO are cylindrical in shape, and we observe similar pattern in their magnetic polarizabilities (M_2 and M_3 are nearly identical and M_1 is typically much larger than M_2 and M_3). We do not observe similar pattern for a rock, which is not metallic and does not have a rotationally symmetric cylindrical shape.

6.3 FEATURE SELECTION

SIG has performed feature extraction on the following data sets collected at camp SLO: EM61-array, Magnetometer, MSEMS, and TEMTADS. Most of these sensors were deployed in survey mode, except TEMTADS. The number of features extracted from some next-generation sensors is very large. For example, TEMTADS collected response at 115 timegates for each of the 625

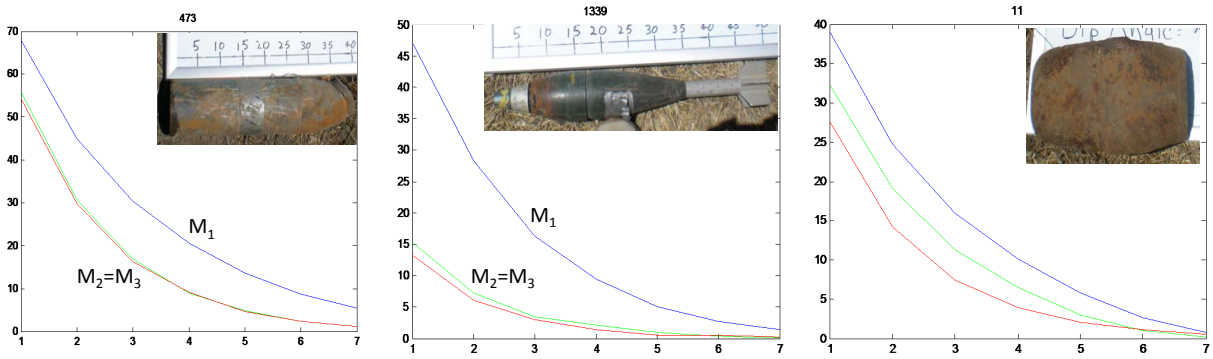


Figure 6: (a) Three principal magnetic polarizabilities of a 4.2-inch mortar (ID:473) is shown as a function of the seven timegates chosen for inversion, (b) Three principal magnetic polarizabilities of a 81-mm mortar (ID:1339) is shown as a function of the seven timegates chosen for inversion. In case of both (a) and (b), we observe M_1 to be significantly larger than other two polarizabilities, and M_2 and M_3 are nearly identical. (c) Three principal magnetic polarizabilities of a clutter (ID:11) is shown as a function of the same seven timegates.

channels (25 transmitters and 25 receivers). Rather than imposing a particular type of decay model for the magnetic polarizabilities extracted from a dipole model, SIG decided to process magnetic polarizabilities for each timegate as individual features. This approach would have led to a very high-dimensional feature space (note that SIG dipole model extracts three principal polarizabilities for each timegate). Classification in a high-dimensional feature space is not robust and generalizable, especially when limited by a small set of training data.

SIG has investigated multiple methods, *e.g.,* relevance vector machine (RVM) [6] and Bayesian elastic net [11], for feature selection and ultimately selected a RVM-based sparse logistic regression model that is linear in feature space. This feature selection approach accomplishes two important goals.

1. Identify Features that are relevant for discriminating UXO from clutter
2. Dimensionality reduction for greater generalizability of the classifier

The sparse logistic regression approach to feature selection assigns a weight to each feature to find the best linear separator in the feature space. Other approaches such as a Forward Selective Search (FSS) cross-validation method finds the best subset of features, but assigns an equal weighting to the selected features, which can be sub-optimal. The sparse logistic method also imposes a Laplacian sparseness prior distribution on the features' relevance to prefer solutions that reduce the overall dimensionality of the space. Dimensionality reduction is particularly important in the current application due to the low number of targets of interest (TOI) in the training data set (29 out of 179 available training objects are TOI and the rest are clutter objects). Keeping the feature dimensionality low ensures that a classifier trained with small set of training data would generalization well on the unseen testing data. The following paragraphs describe the features selected for each of the sensors used by SIG.

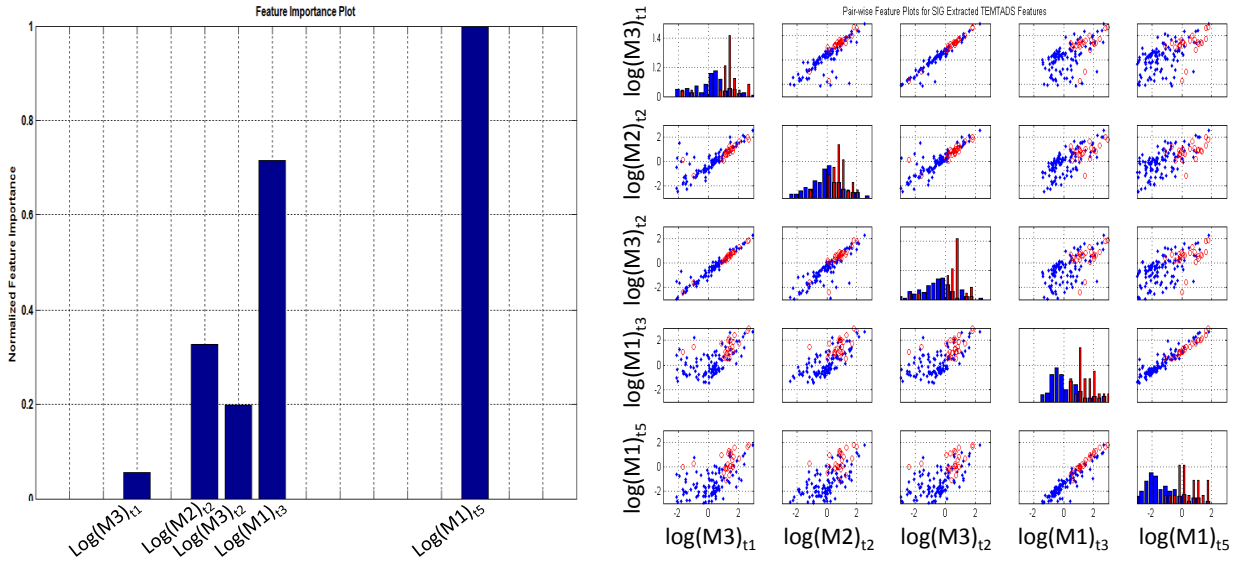


Figure 7: (a) The plot shows the importance of the features in discriminating the labeled training data for TEMTADS. The weights or feature importances are normalized so that most important feature corresponds to unit weight. Note that only five weights are positive, meaning only those features are relevant for classification. The feature weights are evaluated based on 179 labeled training data consisting of 29 UXO and 150 clutter objects. (b) This plot shows pair-wise feature distribution of TOI and clutter data for five relevant features. The off-diagonal plots display feature distribution of both TOI and clutter data collected by TEMTADS, whereas the diagonal plots display one-dimensional distribution along corresponding features for both TOI and clutter (using red and blue histograms). UXO and clutter are linearly separable if the blue and red histogram do not overlap.

TEMTADS is one of the next-generation sensor systems developed specifically for UXO discrimination. Feature selection plays a crucial role for TEMTADS since the size of the parameter vector for the original data would be enormous. After careful analysis, SIG decided to use five logarithmically spaced timegates (at 89, 169, 302, 522 and 890 μsec). SIG extracted three magnetic polarizabilities for each timegate, resulting in a set of fifteen features for each anomaly. As mentioned above, SIG decided to perform dimensionality reduction of the feature space for better generalization. SIG has employed a RVM-based feature selection approach to achieve this task, trained on 179 labeled anomalies.

Figure 7(a) shows the normalized weights or relevance of the features (1 being most important and 0 meaning not important at all). The naming convention for the features is as follows: the largest magnetic moment for timegate n is named as $(\mathbf{M1})_{\mathbf{tn}}$ (where $n = 1, \cdots, 5$), second and third magnetic moments are named M2 and M3, respectively. It is to be noted that the algorithm identifies only five feature vectors to be relevant for classification. The weights (or relevance) of the features were used subsequently to obtain the similarity matrix (the kernel) in the feature space [6]. Since ten out of fifteen features have virtually zero weight, this approach is effectively ignoring those

ten features in designing the classifier. We should also mention here that our feature extraction model could not analyze all anomalies (a total set of 1464 anomalies). There were five anomalies (with IDs 218, 227, 314, 365, and 515) for which SIG could not extract the model parameters reliably due to lack of associated data and/or level of background noise. All five anomalies belong to the testing data set, and SIG included them at the bottom of the TEMTADS prioritized dig list in accordance with the scoring memorandum.

Figure 7(b) shows the pair-wise feature distribution where the targets of interest (TOI) and clutter are represented in red and blue, respectively. For the sake of clarity, we have only shown the distribution for the five features that are deemed important by our feature selection algorithm.

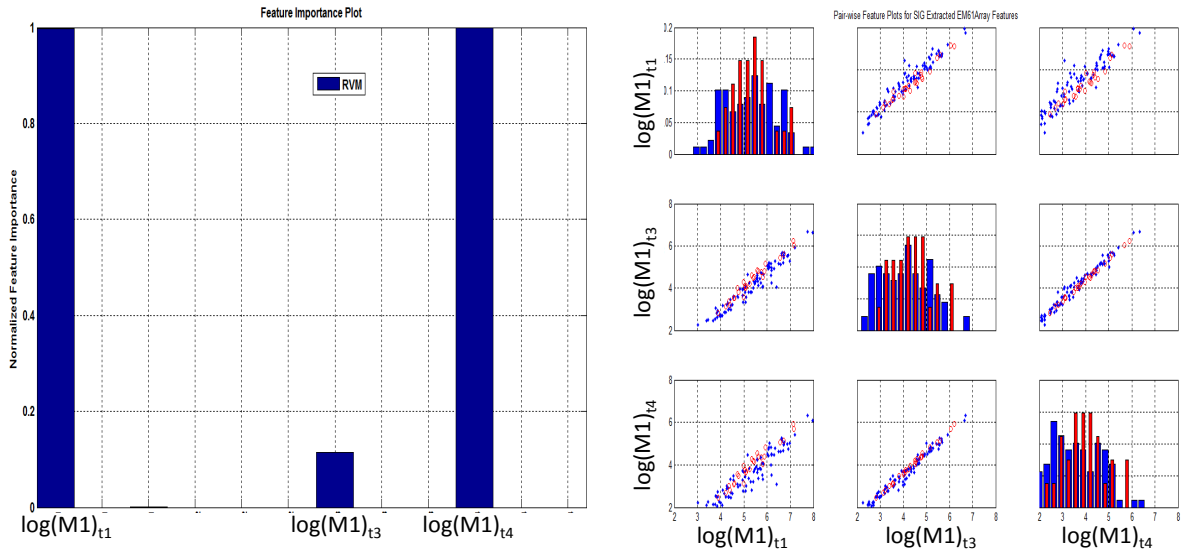


Figure 8: (a) The bar chart shows importance of the features in classifying the labeled training data for EM61-Array. The weights are normalized so that most important feature corresponds to unit weight. Note that only three features are "deemed" relevant for classification. (b) This 3x3 array of plots show pair-wise feature distribution for the EM61-array sensor, displaying TOI and clutter in red and blue, respectively. The diagonal plots show one-dimensional distribution along corresponding features. UXO and clutter are linearly separable if the blue and red histograms do not overlap.

EM61-array deployed during the SLO data collection used four timegates and SIG extracted three magnetic moments (M1, M2 and M3) for each timegate. Based on the available data, SIG could not analyze 38 anomalies due to insufficient data (of which three belongs to the training dataset). Based on the available labeled data for which we could reliably estimate features, we employed the RVM-based linear feature selection algorithm. Figure 8(a) shows the feature importance plot, where only three out of twelve extracted features are deemed relevant for classification. Figure 8(b) shows pair-wise distribution of three important features.

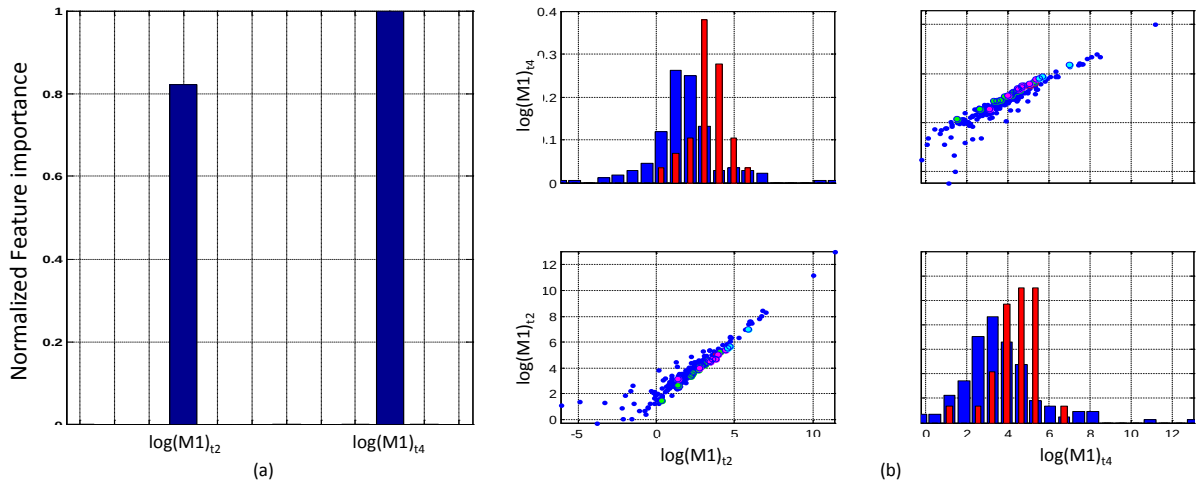


Figure 9: (a) The bar chart shows importance of the features in classifying the training data for dual-mode MSEMS sensor. The plot shows twelve features (three magnetic polarizabilities for each of four timegates) extracted from the EM sensor, while the MAG sensor data has been rejected by SIG on grounds of being very noisy. (b) Pair-wise feature distribution for the MSEMS sensor is shown here for the two selected features.

Figure 9(a) shows the feature selection for MSEMS system. This sensor system incorporated both an EM61 and a MAG sensor. However, SIG decided to ignore the MAG data because they were highly noisy. The EM61 section provided us twelve features (similar to the EM61-array shown above). Feature selection identifies only two features as relevant for UXO discrimination. Figure 9(b) shows the pairwise feature distribution plot for the two chosen features.

Figure 10(a) shows the feature selection for the MetalMapper system. SIG was provided with features for all anomalies detected by the MetalMapper. Original Metalmapper system collected response at fifty timegates, of which first eight timegates were ignored. SIG decided to use features from five timegates uniformly separated within the operational time window. The magnetic polarizabilities corresponding to these timegates are shown in the horizontal axis of Figure 10(a). The feature selection algorithm chooses four features and the corresponding pairwise feature distribution plots are shown in Figure 10(b). One can observe that TOI and clutter features are better separated than most of the other sensors. As a consequence, the classification performance of the MetalMapper was clearly superior to most of the other sensors deployed in the SLO study.

7 PERFORMANCE ASSESSMENT

SIG could not analyze five out of 1464 raw anomaly responses for the TEMTADS system, which is much lower than SIG's proposed success criteria of reliable feature extraction of 90% of the anomalies. Similarly, SIG could not analyze 38 out of 1464 EM61-array responses and 9 out of 1498 MAG responses. Note that in both cases SIG had extracted features for more than 90% of the anomalies.

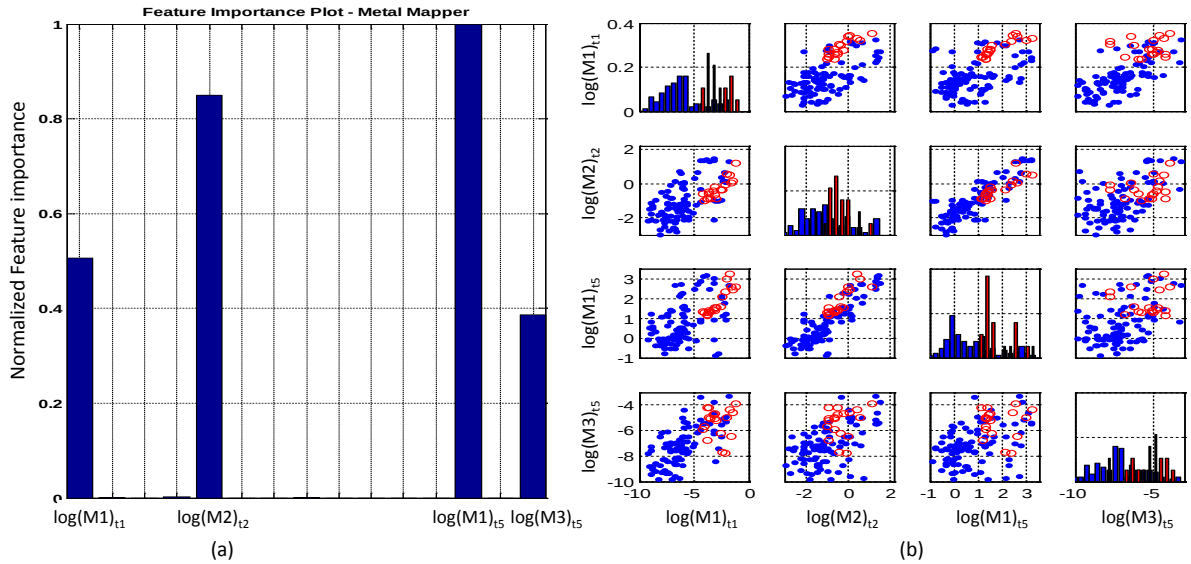


Figure 10: (a) The bar chart shows importance of the features in classifying the MetalMapper training data. The plot shows fifteen features (three magnetic polarizabilities for each of five timegates) extracted from the sensor. (b) Pair-wise feature distribution for the MetalMapper sensor is shown here for the four selected features.

Once the features have been reliably extracted from the labeled data, we focus on training a classifier that is capable of discriminating UXO from clutter. As discussed in the demonstration plan [1], SIG proposed to train two classifiers for each sensor-extractor pair. The supervised classifier uses only labeled features for training, whereas the semi-supervised classifier uses both labeled and unlabeled data. SIG has employed the same Relevance Vector Machine (RVM) [6] based supervised classifier as employed during the Sibert data analysis.

7.1 PERFORMANCE ANALYSIS OF PASSIVE CLASSIFIERS

This subsection will analyze the performance of passive classifiers (supervised and semi-supervised), where the labeled training data is provided. The supervised classifier does not account for the context provided by the unlabeled data. In contrast, a semi-supervised approach, implemented via a graph-based classification scheme [7] is capable of exploiting the unlabeled data in addition to the labeled data when building the classifier. The classifier, titled PNBC [7], uses random Markov walk to define the neighborhood of each feature vector and define the similarity between adjacent feature vectors. The algorithm is based on the presumption that feature vectors that are similar in feature space are likely to share the same label (UXO or non-UXO, for a binary classification). We have implemented a linear PNBC classifier in this case where the final classifier boundary is a linear hyperplane in feature space. In the following paragraphs, we shall analyze the performance of both classifiers on each sensor data. For some sensors, SIG had features from two sources (internal feature extraction and features provided by UBC). We shall compare their classification performances using the receiver operating characteristics (ROC) curves.

7.1.1 SETTING OF CLASSIFICATION THRESHOLDS

The supervised and semi-supervised classifier are trained on data provided by the ESTCP program office. Once trained, each of the classifiers yields a statistical estimate (or a likelihood) of the label of a given feature vector from the testing data. As mentioned in the scoring memorandum, SIG was supposed to produce a prioritized dig list for each classifier where anomalies are ranked in decreasing likelihood of being a clutter (*i.e.,* most likely clutter would be at the top of the list and most likely TOI would be at the bottom). In addition, SIG had the task of assigning a threshold for each dig list such that all test anomalies above the threshold are considered as clutter by SIG with high confidence. SIG has followed a principled Bayesian risk analysis to estimate this threshold, one for each classifier trained on a feature set.

Let x represent a given feature vector under test, and our goal is to estimate the label l, where $l = 1$ is chosen to correspond to a UXO and $l = 0$ to a clutter. Our algorithm yields the probability $p(l = 1|x)$, and $p(l = 0|x) = 1 - p(l = 1|x)$. Let the cost of declaring an item to be a UXO when it is actually a non-UXO be denoted C_{10}, while the cost of declaring an item non-UXO when it is actually a UXO is denoted C_{01}. We set the cost (reward) associated with making a correct classification to zero: $C_{11} = C_{00} = 0$. Given a feature vector x under test, the expected cost (or risk) of declaring the associated item to be a UXO is

$$R_{UXO} = C_{10} p(l = 0|x) \tag{21}$$

while the risk of declaring the item to be non-UXO is

$$R_{non-UXO} = C_{01} p(l = 1|x) \tag{22}$$

Our objective is to minimize the risk, and therefore we declare x to be a UXO if $R_{UXO} < R_{non-UXO}$, and otherwise we declare as non-UXO. Hence, we declare the item to be a non-UXO if

$$\frac{p(l = 0|x)}{p(l = 1|x)} > \frac{C_{01}}{C_{10}}. \tag{23}$$

Thus, by selecting the costs C_{01} and C_{10}, and given a statistical measure $p(l = 1|x)$, one defines the threshold or operating point on the ROC. Note that the more the relative cost of a missed UXO increases, the greater the ratio $p(l = 0|x)/p(l = 1|x)$ to be to leave an item unexcavated. Therefore, in setting our threshold, we will set this ratio $C = \frac{C_{01}}{C_{10}}$. By varying the threshold C, one maps out the receiver operating characteristic (ROC). Rather than setting a fixed threshold C (as done during the Ford Ord analysis), SIG decided to perform complete risk analysis by varying C over a range and obtain the corresponding detection probabilities (Pd, evaluated only on the labeled data). The three step procedure for setting threshold for each feature set is shown in Fig. 11. First, we train a classifier on labeled data (including unlabeled data for semi-supervised classifier) and generate a ROC curve using Leave-one-Out cross-validation (shown in Fig. 11(a) for RVM-based supervised classifier trained on labeled TEMTADS data). Figure 11(b) shows the variation of the probability of detection Pd (evaluated only on labeled data) and percentage excavation as a function of cost of miss, C. Note that the classifiers detects all UXO (Pd=1) corresponding to $C = 45$. The plot

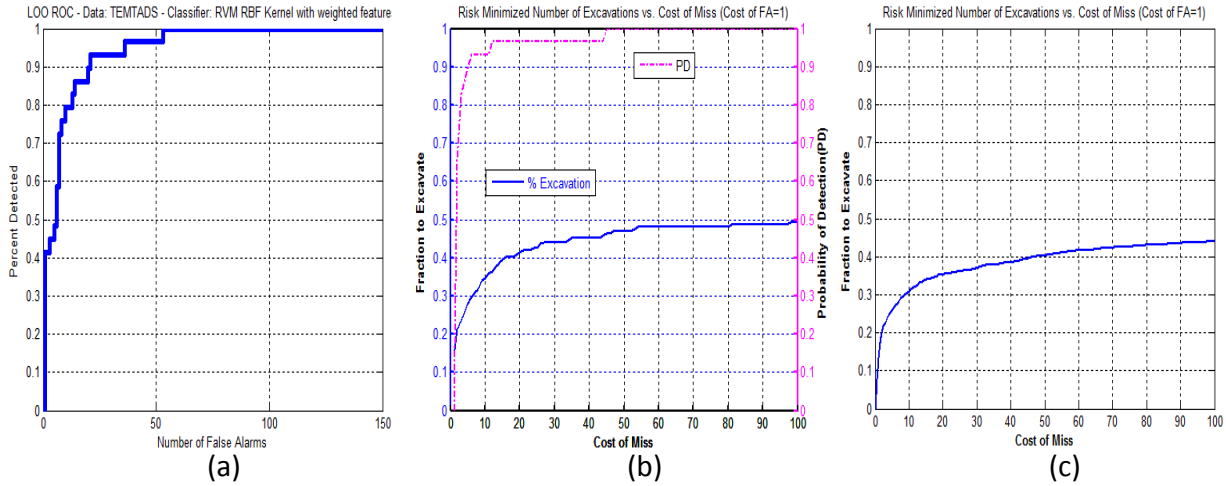
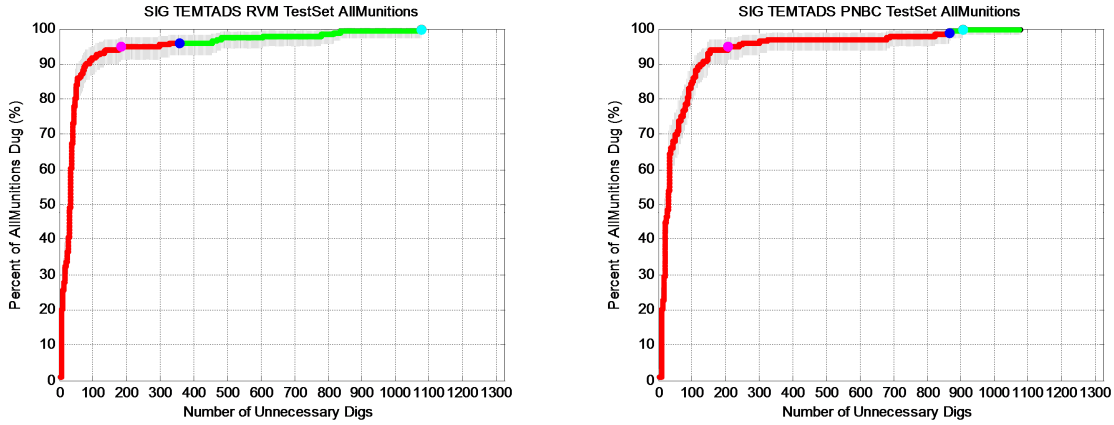


Figure 11: (a) This plot shows Receiver Operating Characteristics of RVM-based supervised classifier trained on labeled TEMTADS data. The percentage of UXO detected (Pd) is plotted as a function of the number of false alarms (Nfa) generated by the classifier. (b) This plot show the Risk Analysis, where the probability of detection (in red) and percentage excavation (in blue) are shown as a function of the relative cost of miss. The probability of detection achieves 100% at $C = 45$. SIG chose $C = 50$ for the testing data. (c) This plot shows the percentage of excavation for the testing data as a function of the threshold C. Note that SIG would propose to excavate 40% of the buried anomalies once it sets the threshold to 50.

also shows that one needs to excavate 48% of the data in order to detect all UXO. Although this analysis is based solely on the limited training data set (179 labeled samples containing only 29 UXO), the generalizability of the trained classifiers allowed SIG to use the same threshold value for the TEMTADS supervised classifier. The same procedure is followed for all sets of features collected from different sensors.

7.1.2 TEMTADS CLASSIFIER PERFORMANCE

SIG has performed both feature extraction and classification for the TEMTADS data. In accordance with the SIG Demonstration plan [1], SIG delivered two prioritized dig list for TEMTADS, one corresponding to the supervised classifier, and the other for the semi-supervised one. The performance of the classifiers were analyzed by IDA. Figure 12(a) shows the performance of the supervised classifier. Note that the blue circle corresponds to the threshold chosen by SIG, the magenta circle corresponds to 95% probability of detection, and the cyan circle correspond to Pd=1 with least number of unnecessary digs (false alarms). According to the ROC plot, SIG would have left approximately 3% of the UXO (missing an UXO) in the ground. Figure 12(b) shows the performance of the semi-supervised classifier. Note that SIG's choice of the threshold (blue circle) was very close the ideal threshold value (cyan circle) (remember that one needs to know all testing labels to generate the cyan circle and SIG was provided with labels for only training data). Both of the classifiers obtained above 90% UXO detection with very low false alarm rate (< 100 for supervised and < 200 for semi-supervised classifier).



(a) (b)

Figure 12: (a) This plot shows supervised classifier performance on TEMTADS features extracted by SIG, (b) This plot shows semi-supervised classifier performance on TEMTADS features extracted by SIG. The blue circle on both plots represent the threshold chosen by SIG, whereas the cyan circle corresponds to the ideal threshold choice (corresponding to Pd=1), and the magenta circle corresponds to Pd=0.95.

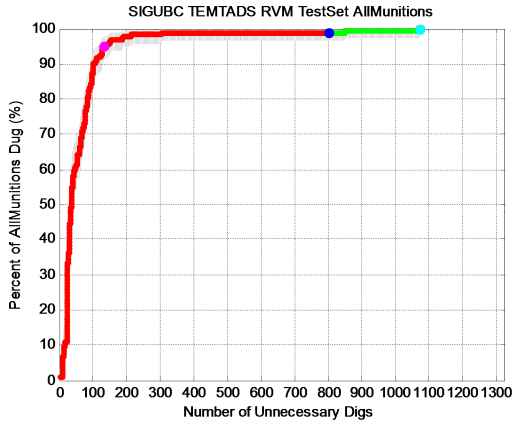
In addition to the TEMTADS features extracted internally, SIG was also provided with features extracted by UBC/Sky Research. Since UBC extracted features for all 115 timegates, SIG has performed feature selection and subjected only the relevant features to classifier training.

Figure 13(a) shows the supervised classifier performance on UBC-extracted TEMTADS features. According to the ROC plot, UBC-chosen threshold (the blue circle) obtained near-perfect UXO detection, although the ideal threshold choice would have excavated almost all buried anomalies (since the cyan circle is at the top-right corner of the ROC plot). Figure 13(b) shows the performance of the semi-supervised classifier, which is similar to its supervised counterpart. In comparison, the SIG semi-supervised classifier would have left almost 200 anomaly on the ground.

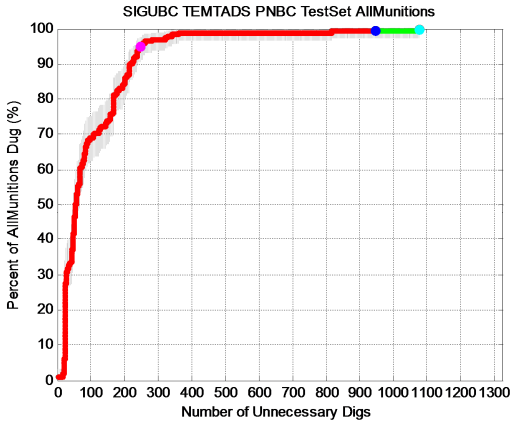
7.1.3 METALMAPPER CLASSIFICATION PERFORMANCE

SIG did not perform feature extraction for the MetalMapper data. Provided with magnetic moment features for all available timegates, SIG decided to choose five timegates spread evenly over the active time window. The main reason to prune manually rather than allowing the feature selection algorithm to select relevant features from all timegates was the lack of training data for robust estimation of feature importance. The manually subsampled feature set was subjected to feature selection, before training the supervised and semi-supervised classifiers.

Figure 14(a) shows the supervised classifier performance on relevant MetalMapper features. The performance has been shown to be clearly superior to any other supervised classification results. The classifier obtained 95% detection with less than 50 false alarms and detected all UXO with only 600 false alarms, leaving more than 400 anomalies on the ground. Figure 14(b) shows the performance of the semi-supervised classifier on MetalMapper data. This performance has been

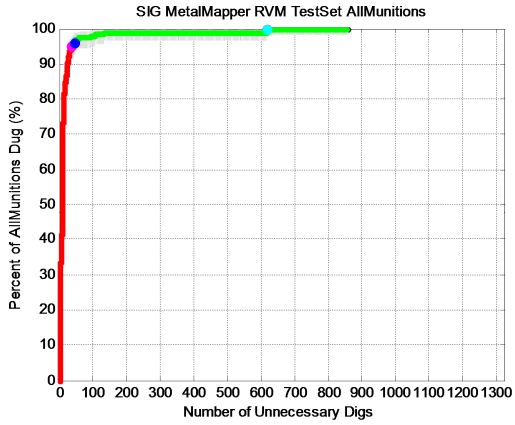


(a)

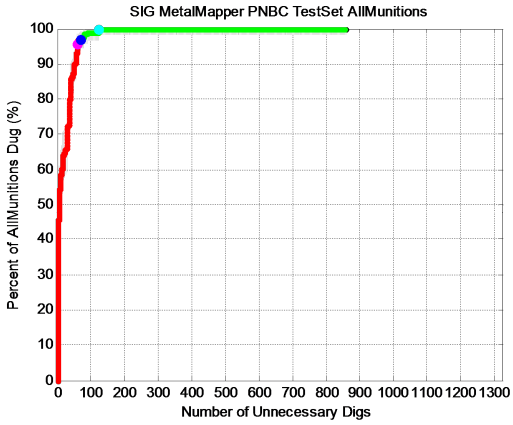


(b)

Figure 13: (a) This plot shows supervised classifier performance on TEMTADS features extracted by UBC, (b) This plot shows semi-supervised classifier performance on TEMTADS features extracted by UBC. The blue circle on both plots represent the threshold chosen by SIG, whereas the cyan circle corresponds to the ideal threshold choice (corresponding to Pd=1), and the magenta circle corresponds to Pd=0.95.



(a)



(b)

Figure 14: (a) This plot shows supervised classifier performance on MetalMapper data. The features were provided by Snyder Geoscience, which were further pruned by SIG feature selection algorithm. (b) This plot shows semi-supervised classifier performance on relevant MetalMapper features. The classifier detects all TOI with only 120 false alarms.

proclaimed by ESTCP as the best performance achieved by any sensor-classifier combination at SLO. The classifier detects all UXO with only 120 false alarms. Note that MetalMapper classification performance has been much superior to the proposed criteria of rejecting more than 30% of the clutter while detecting all UXO objects.

7.1.4 BUD CLASSIFICATION PERFORMANCE

SIG did not perform feature extraction for the data collected by the Berkeley UXO Discriminator (BUD). Provided with magnetic moment features for all available timegates, SIG decided to choose five timegates spread evenly over the active time window. Figure 15(a) and(b) show performance on BUD features using supervised and semisupervised classifier, respectively. Note that BUD classification performance did not satisfy the proposed criteria of rejecting more than 30% of the clutter while detecting all UXO objects. The main reason for the poor performance of BUD seems to be coming from poor feature extraction and insufficient collection of representative clutter samples for a robust training of the classifiers.

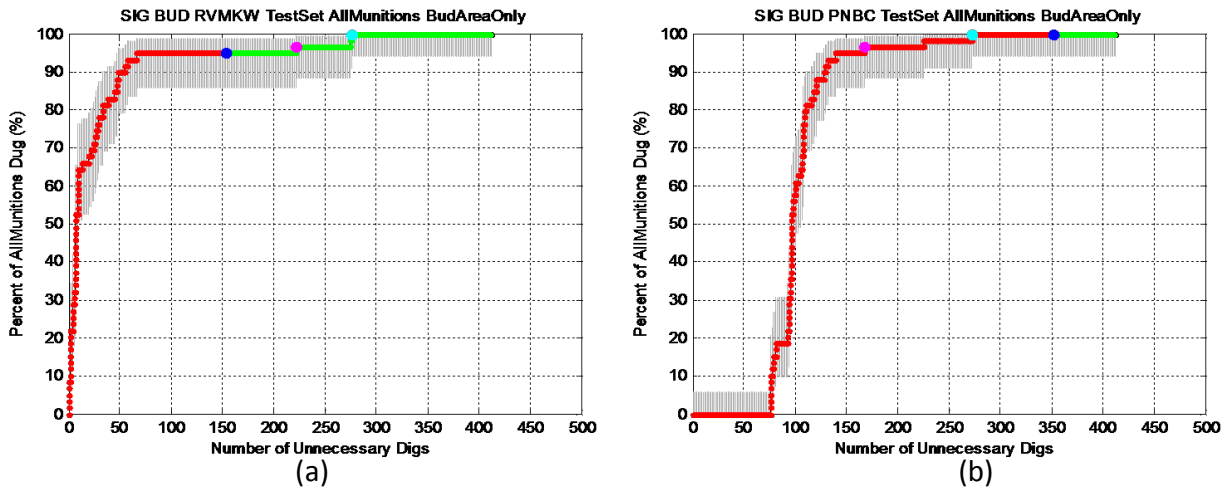
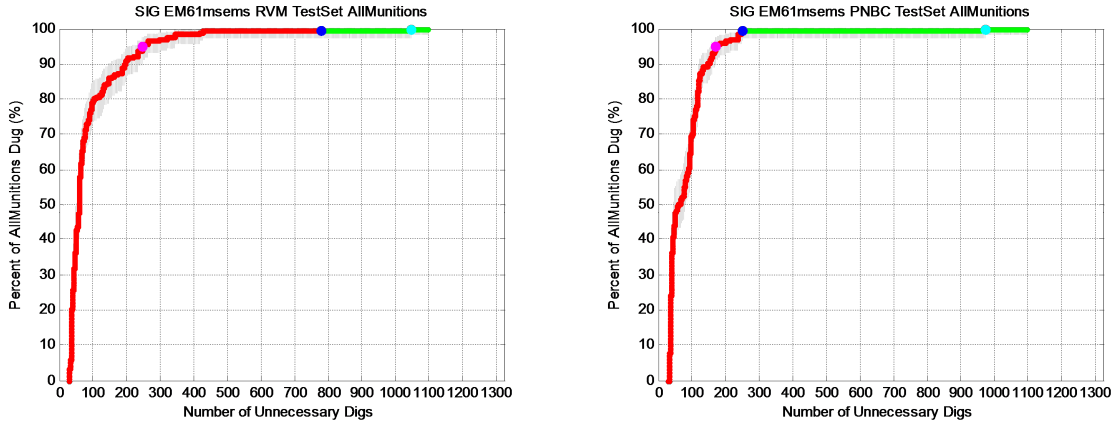


Figure 15: (a) This plot shows supervised classifier performance on BUD data. The features were provided by LBL. (b) This plot shows semi-supervised classifier performance on relevant BUD features.

7.1.5 MSEMS CLASSIFICATION PERFORMANCE

MSEMS is a dual-mode sensor system that incorporated a EM61 and a MAG sensor. SIG has performed its own feature extraction for MSEMS, along with features provided by UBC. Careful preprocessing of the raw data indicated that MAG response is highly noisy. Therefore, SIG proceeded to extract features only for the EM61 section. Figure 16(a) and (b) show the supervised and semi-supervised classifier performances on MSEMS features extracted internally at SIG. The performance of the semi-supervised classifier is significantly better the supervised classifier. However, none of the two classifiers could reject 30% of the clutter while detecting all targets-of-interest. The main reason being a single TOI (2.36") that was broken into multiple pieces, making it very difficult to identify. If this anomaly is ignored, the semi-supervised classifier obtains 100% TOI discrimination while rejecting more than 70% of the clutter items.

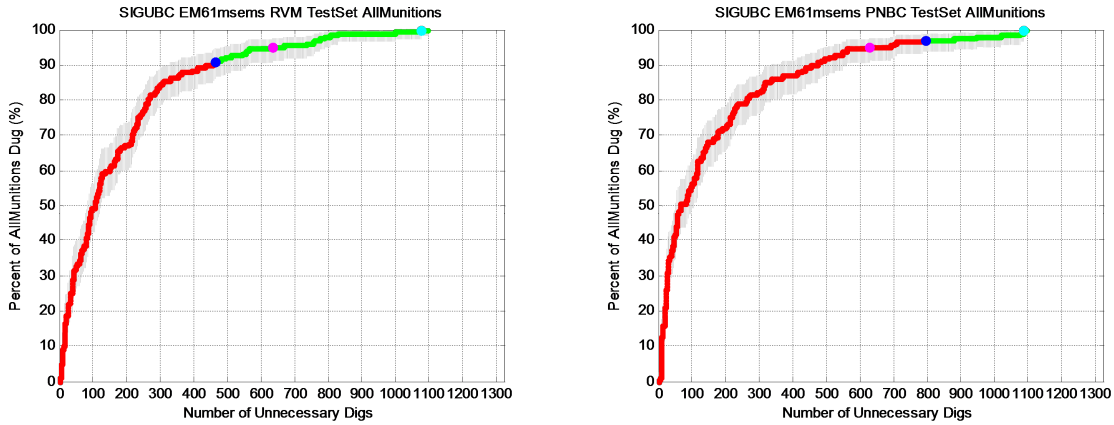
SIG has also trained supervised and semi-supervised classifier on MSEMS features extracted by UBC/Sky Research. Figure 17(a) and (b) show the supervised and semi-supervised classifier performances on MSEMS features extracted by UBC. The performance of both classifiers are quite



(a) (b)

Figure 16: (a) This plot shows supervised classifier performance on SIG-extracted features obtained from raw MSEMS data. MAG data belonging to MSEMS is ignored for feature selection and classification. (b) This plot shows semi-supervised classifier performance on SIG-extracted MSEMS features. Performance of the semi-supervised classifier is much superior to its supervised counterpart.

similar and are both inferior to the semi-supervised classification performance on SIG-extracted features (shown in Fig. 16(b)).



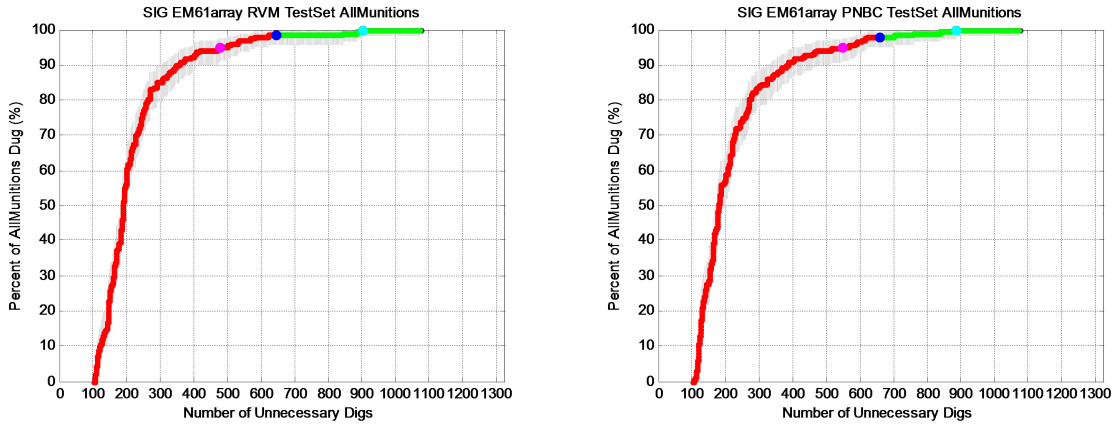
(a) (b)

Figure 17: (a) This plot shows supervised classifier performance on UBC-extracted features obtained from raw MSEMS data, (b) This plot shows semi-supervised classifier performance on UBC-extracted MSEMS features.

7.1.6 EM61-ARRAY CLASSIFICATION PERFORMANCE

EM61 is a time-domain electromagnetic sensor (TEM) that is being used for UXO detection and discrimination for almost a decade. In accordance with the demonstration plan, SIG has performed

its own feature extraction for EM61-array, along with features provided by UBC. EM61-array used four timegates resulting in twelve magnetic moment features. Feature selection reduced the dimensionality of the feature space to two relevant features. Figure 18(a) and (b) show the supervised and semi-supervised classifier performances on EM61-array features extracted internally at SIG. The performance of both classifiers are similar.



(a) (b)

Figure 18: (a) This plot shows supervised classifier performance on SIG-extracted features obtained from raw EM61-array data, (b) This plot shows semi-supervised classifier performance on SIG-extracted EM61-array features.

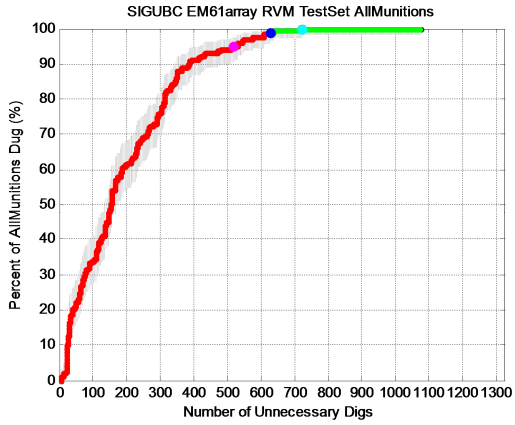
SIG has also trained supervised and semi-supervised classifier on EM61-array features extracted by UBC/Sky Research. Figure 19(a) and (b) show the supervised and semi-supervised classifier performances on EM61-array features extracted by UBC. The performance of both classifiers are quite similar. The main noticeable difference between the SIG and UBC feature performances is that ROC plots for UBC start at origin (Pd=0,Pfa=0), whereas the SIG ROC plot start at (Pd=0,nfa=100). However, both sets of ROC plots catch up as Pd reaches one.

7.1.7 EM61-CART CLASSIFICATION PERFORMANCE

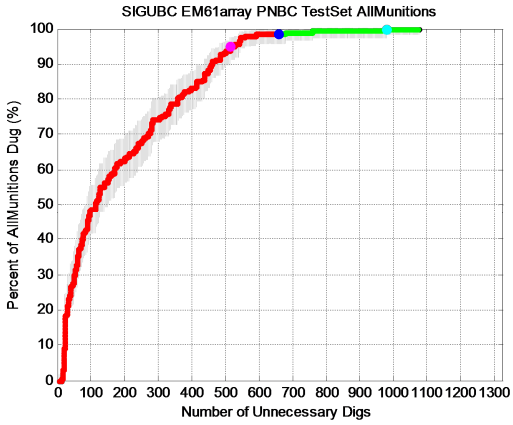
SIG did not perform feature extraction for EM61-cart data. Features extracted by UBC/Sky Research were provided to SIG for subsequent feature selection and classifier training. Figure 20(a) and (b) show the supervised and semi-supervised classifier performances on EM61-cart features. The performance of both classifiers are quite similar to each other.

7.1.8 MAG CLASSIFICATION PERFORMANCE

Magnetometer provided a single-dimensional features (strength of the magnetic field). Hence no feature selection step was necessary. SIG has performed supervised and semi-supervised classification using MAG feature extracted internally and externally (provided by UBC/Sky research). Figure 21(a) and (b) show the supervised and semi-supervised classifier performances on SIG features, and Fig. 22(a) and (b) show the supervised and semi-supervised classifier performances on

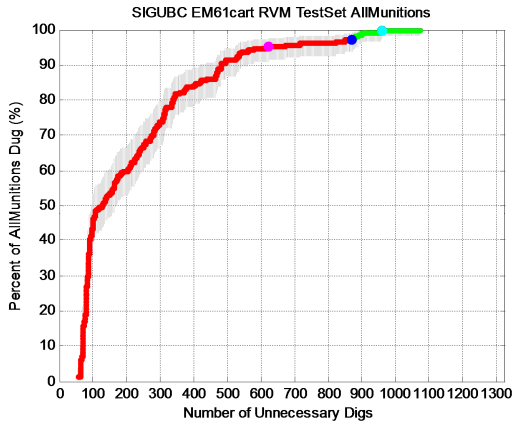


(a)

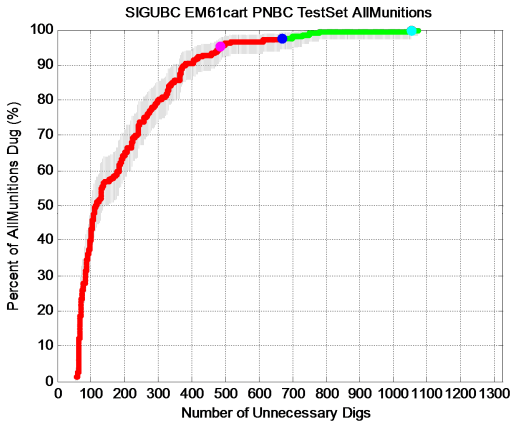


(b)

Figure 19: (a) This plot shows supervised classifier performance on UBC-extracted features obtained from raw EM61-array data, (b) This plot shows semi-supervised classifier performance on UBC-extracted EM61-array features.



(a)



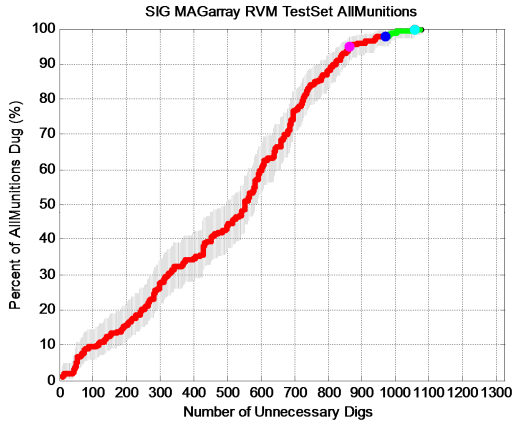
(b)

Figure 20: (a) This plot shows supervised classifier performance on UBC-extracted features obtained from EM61-cart data, (b) This plot shows semi-supervised classifier performance on UBC-extracted EM61-cart features.

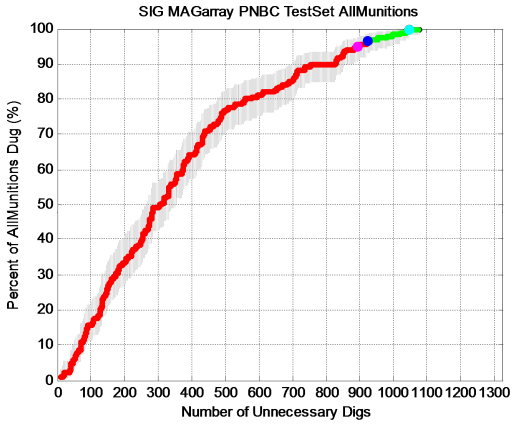
UBC features. The performance of both sets of classifiers are similar, but their performances were much inferior to next-generation sensors, such as MetalMapper and TEMTADS.

7.2 ACTIVE LEARNING CLASSIFICATION PERFORMANCE

SIG has developed active learning classifier for both TEMTADS and EM61-array data. The algorithm, developed at Duke [12], was also used in the Sibert study. Hence, we shall only highlight the salient features of the active learning algorithm here. Given that all the anomalies will be eventually excavated for the SLO demonstration, one may ask in which order these anomalies should

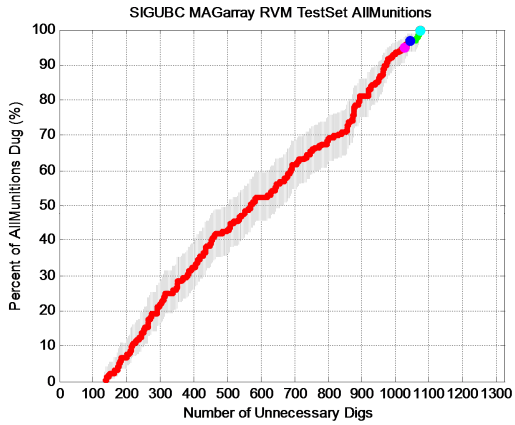


(a)

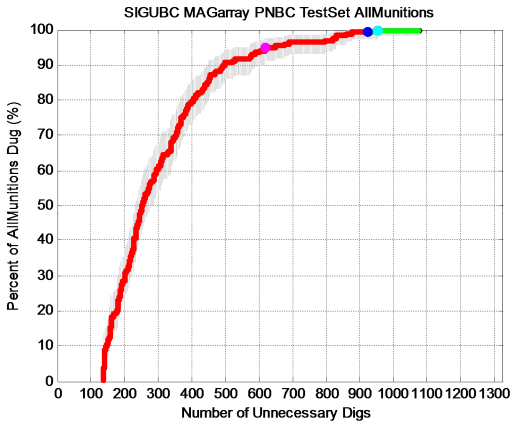


(b)

Figure 21: (a) This plot shows supervised classifier performance on SIG-extracted MAG features, (b) This plot shows semi-supervised classifier performance on SIG-extracted MAG features.



(a)



(b)

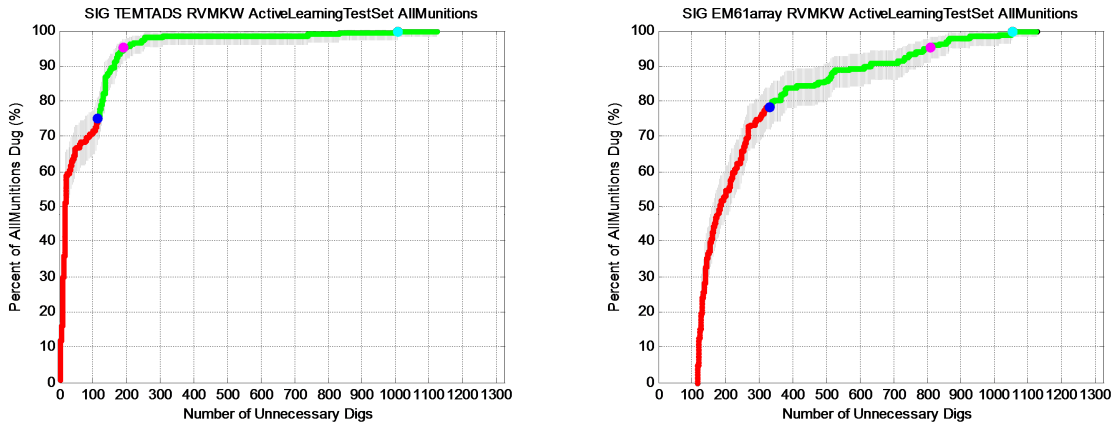
Figure 22: (a) This plot shows supervised classifier performance on UBC-extracted MAG features, (b) This plot shows semi-supervised classifier performance on UBC-extracted MAG features.

be excavated that improves the performance of the classifier algorithm most. It has been shown that [9] this question can be answered in a quantitative information-theoretic manner, where the objective would be to choose the feature vector for labeling that maximizes the mutual information between the classifier and the new data point to be labeled.

The active learning algorithm starts with the selection of a set of representative samples from the available data. The samples, known as basis functions, are chosen by the active learning algorithm to obtain a broad representation of the data manifold. Labels of the basis vectors (corresponding anomalies) are requested as an initial starting point for the classifier. Note that, the selection step stops when the information gain is lower than a predefined threshold. Once the classifier is initialized, at each step it will identify a single buried anomaly that would be most informative for

the classifier to know the label, and the label would be obtained through digging. In reality all the SLO anomalies were dug up and the ESTCP Program Office had labels of all the anomalies. The active learning algorithm proceeds by choosing the most informative sample at every iteration, followed by obtaining the corresponding label from ESTCP Program Office. Once the label is obtained, the classifier is retrained and the next most informative sample is identified. This iterative sample selection, labeling, and retraining step involves multiple communication between SIG and ESTCP Program Office. In order to reduce the communication overhead, SIG has employed a simplified form, where the algorithm identified a set (typically 10) of informative samples it needs to be labeled. Subsequently the algorithm was retrained based on the labels provided by ESTCP.

Figure 23(a) displays the performance of an active-learning classifier trained with 120 labeled samples from TEMTADS dataset chosen adaptively by our algorithm. In contrast, we obtain similar classification performance using a fully supervised classifier trained with 179 labeled samples as provided by ESTCP Program Office. Adaptive selection of labeled training data is more efficient in obtaining optimal classification performance in this case. Similarly, Fig. 23(b) displays the performance of an active-learning classifier trained with 133 labeled samples from the EM61-array dataset chosen adaptively by our algorithm.



(a) (b)

Figure 23: (a) This plot shows the performance of the actively learned classifier trained with adaptively selected 120 labeled samples from TEMTADS dataset, (b) This plot shows the performance of the actively learned classifier trained with adaptively selected 133 labeled samples from EM61-array dataset.

8 COST ASSESSMENT

SIG took part in developing physics-based models for buried UXO and employing passive (supervised and semi-supervised) and active classifiers for discriminating UXO from clutter. The demonstration was conducted with the support of several SIG personnel. Dr. Lawrence Carin (PI) acted as the Quality Assurance (QA) Officer, and also managed the demonstration process and reporting. Dr. Xianyang Zhu, Mr. Levi Kennedy, and Dr. Nilanjan Dasgupta performed the data

processing and analysis. Dr. Paul Runkle provided cost management and general oversight. The anticipated cost basis for this technology demonstration is the number of anomalies that can be classified per unit time, while taking into account classifier training time. The main cost driver of this technology is expected to be time it takes to perform the dipole model inversion. As a result, the cost will be a function of the number of anomalies, and hence also a function of the site size. The cost table structure for the proposed work and the budget are shown in Figure 24.

There is additional cost for developing two sets of passive classifiers (supervised and semi-supervised) for each sensor-extractor combinations. Note that SIG did not perform feature extraction for BUD and MetalMapper, although it has performed feature selection and subsequent classifier design for all sensor-extractor pairs. The SLO demonstration produced 1464 raw anomaly responses for TEMTADS, EM61-array, 1479 responses for EM61-cart and 539 responses for BUD sensor. The cost is approximately linearly proportional to the number of feature vectors extracted, and hence the overall cost should be scaled appropriately for a new site, based on the number of anomalies detected at that site.

	Features				Algorithm			Datasets Analyzed	$K	
	SIG	UBC	LBL	SNYDER	Sup	Semi-Sup	Active Learning			
Magnetometer	X	X			X	X		4	29.23	
MetalMapper				X	X	X		2	14.62	
EM61-Cart		X			X	X		2	14.62	
EM61-Array	X	X			X	X	X	6	43.84	
TEMTADS	X	X			X	X	X	6	43.84	
BUD			X		X	X		2	14.62	
MSEMS	X	X			X	X		4	29.23	
								26	190	Total

Figure 24: Itemized Cost for the SLO Data Analysis

9 IMPLEMENTATION ISSUES

9.1 ENVIRONMENTAL CHECKLIST

Not applicable to this demonstration.

9.2 END-USER ISSUES

Not applicable to this demonstration.

References

[1] L. Carin, "ESTCP UXO Demonstration Study - Demonstration Plan," *accepted by ESTCP Program Office*, May 2009.

[2] L. Kennedy, "Signal Innovations Group, Inc. San Louis Obispo (SLO) Discrimination Study Training Memo," *accepted by ESTCP Program Office*, September 2009.

[3] B. Delaney and D. Etter, *Report of the Defense Science Board Task Force on Unexploded Ordnance*, http://www.acq.osd.mil/dsb/reports/uxo.pdf, December 2003.

[4] X. Z. Y. Y. Lawrence Carin, Levi Kennedy and D. Williams, "Final Report on SIG Analysis of Sibert Data," *accepted by ESTCP Program Office*, March 2008.

[5] ESTCP, "2009 ESTCP UXO CLASSIFICATION STUDY, SAN LUIS OBISPO, CA," March 2009.

[6] M. E. Tipping, "Sparse bayesian learning and the relevance vector machine," *Journal of Machine Learning Research*, pp. 211–244, 2001.

[7] Q. Liu, X. Liao, and L. Carin, "Detection of Unexploded Ordnance via Efficient Semisupervised and Active Learning," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, no. 9, pp. 2558–2567, September 2008.

[8] Y. Zhang, L. Collins, H. Yu, C. Baum, and L. Carin, "Sensing of unexploded ordnance with magnetometer and induction data: Theory and signal processing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, p. 10051015, November 2003.

[9] V. V. Fedorov, *Theory of Optimal Experiments*. Academic Press.

[10] Parsons, Inc., "Final site inspection report, former camp san luis obispo, san luis obispo, ca," September 2007.

[11] M. Chen, D. Carlson, A. Zaas, C. Woods, G. S. Ginsburg, A. H. III, J. Lucas, and L. Carin, "The Bayesian Elastic Net: Classifying Multi-Task Gene-Expression Data," *submitted to Journal of Machine Learning Research*, August 2009.

[12] Y. Zhang, X. Liao, and L. Carin, "Detection of buried targets via active selection of labeled data: application to sensing subsurface uxo," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, pp. 2535–2543, 2004.

Point of Contact

Point of Contact (Name and Address)	Organization (Name and Address)	Phone Fax Email	Role in Project
Herb Nelson	ESTCP 901 N. Stuart St. Suite 303 Arlington, VA 22203	Tel: 703-696-8726 Fax: 703-696-2114 Herbert.Nelson@osd.mil	Program Manager
Lawrence Carin	SIG 1009 Slater Rd. Suite 200 Durham, NC 27703	Tel:919-660-5270 Fax: 919-660-5293 lcarin@siginnovations.com	Principal Investigator
Levi Kennedy	SIG	Tel: 919-323-3456 Fax: 919-287-2578 lkennedy@siginnovations.com	Project Manager
Xianyang Zhu	SIG	Tel:919-323-3455 xianyang@siginnovations.com	Lead Scientist
Nilanjan Dasgupta	SIG	Tel: 919-475-1103 nilanjan@siginnovations.com	Scientist