INTERIM REPORT

Decision Support Tools for Munitions Response Performance Prediction and Risk Assessment

SERDP Project MR-2226

JANUARY 2013

Laurens Beran
Black Tusk Geophysics Inc.

Stephen Billings
The University of British Columbia

This document has been cleared for public release

SERDP
DOD ▪ EPA ▪ DOE

This report was prepared under contract to the Department of Defense Strategic Environmental Research and Development Program (SERDP). The publication of this report does not indicate endorsement by the Department of Defense, nor should the contents be construed as reflecting the official policy or position of the Department of Defense. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the Department of Defense.

Executive Summary

This report summarizes progress for the first two quarters of SERDP project MR-2226. The primary objective of the initial phase of the project is the development of a decision support system (DSS) to predict classification performance given environmental variables. This report details progress on the following tasks:

(1) *Prediction of target detection thresholds.* Current methods for setting target detection thresholds are based on the worst case (i.e. minimum) amplitude response for a given target of interest. The geophysical system verification (GSV) calculator previously developed by ESTCP assumes that this worst case corresponds to a horizontal target oriented cross-track. We show that, depending on sensor geometry and offset from the center of the array, the minimum target response for a horizontal target is not always in a cross track azimuthal orientation. In addition, we show that the worst case target amplitude can, in practice, produce a prohibitively large number of target picks. To reduce the number of picks, we simulate the distribution of target response amplitudes assuming a uniform distribution of target azimuth and location across the sensor swath. We then show that setting a threshold that provides a very high confidence of detecting targets of interest (e.g. 99 %) can drastically reduce the number of target picks.

(2) *Noise estimation.* Most synthetic analyses use discrete background measurements to characterize noise at a site. However, we find that a simple noise model derived solely from background measurements (e.g. independent, Gaussian noise) cannot accurately reproduce the distributions of dipole polarizabilities recovered from field data. We instead use the recovered polarizability distribution of industry standard objects (ISOs) as a barometer for characterizing noise across a site. We develop techniques to invert the covariance of polarizabilities within the ISO class to recover an effective covariance of the noise on the data across the site. The resulting noise is highly correlated and only approximately Gaussian.

(3) *Performance prediction.* Our aim here is to predict the receiver operating characteristic (ROC) that would be obtained for a sensor and targets under specified environmental conditions. For a given site, we associate the noise covariance derived from ISO polarizabilities with the median location of ISO items, relative to the sensor. With this noise covariance, we can predict the covariance of target polarizabilities at any other location via a straightforward linear analysis.

Next, we derive analytic expressions for the misfit of log-transformed polarizabilities with respect to reference, or library, polarizabilities. This misfit is the statistic most commonly used to discriminate between targets of interest (TOI) and non-TOI using cued sensor data. The distribution of the decision statistic then maps to the predicted ROC. This machinery allows us to efficiently predict how the spatial distribution of targets affects classification performance. We derive analytic expressions for the covariance of target polarizabilities under a uniform distribution of target orientations. We show that target dip has an important effect on classification performance, with a uniform distribution of dips producing ROC performance that is intermediate to horizontal and vertical cases.

Finally, we extend our performance prediction methods to dynamic sensors and develop techniques to account for the effect of varying data density on classification performance. We derive expressions for the distribution of the polarizability decay parameter, which is typically used to classify data acquired with the EM-61 sensor. This provides a means to predict the baseline ROC for a site, and so to predict whether cued interrogation will produce a significant improvement in classification performance.

Contents

Executive Summary 2
List of Figures 2
List of Tables 3
List of Acronyms 4
1. Introduction 5
2. Background 5
2.1. Data acquisition 5
2.2. Feature extraction 6
2.3. Classification 9
3. Predicting target response thresholds 11
4. Performance prediction 16
4.1. Performance prediction for cued interrogation 17
4.1.1. Noise estimation 17
4.1.2. Predicting the distributions of polarizabilities 27
4.1.3. Predicting the ROC 35
4.2. Performance prediction for dynamic sensors 42
5. Conclusions and further work 48
References 49
Appendix A. Summary of ESTCP demonstration data sets 51
Appendix B. Expected polarizability covariance under a uniform distribution of target orientations 53

List of Figures

1 Electromagnetic induction (EMI) survey 5
2 Sequential inversion approach 9
3 Computing the ROC from the score distributions of TOI and non-TOI 11
4 Anomaly amplitude analysis for a three coil EM-61 array 12
5 Anomaly amplitude analysis for an asymmetric three coil EM-61 array. 13
6 Cumulative distribution of response amplitudes 14
7 Cumulative distribution of response amplitudes at JDR 15
8 Distribution of response amplitudes for targets distributed uniformly in azimuth and dip. 15
9 Methods for polarizability prediction 17
10 Background noise standard deviations for MetalMapper sensor at Camp Butner and Pole Mountain. 18
11 Components of the observed secondary field for MetalMapper dynamic data acquired at Camp Butner 19
12 Estimated ISO polarizabilities at Camp Beale and Pole Mountain. 19
13 Synthetic example of estimation of effective noise standard deviation 21
14 Estimates of effective noise standard deviations for MetalMapper Beale Parsons and Pole Mountain data sets 22
15 Comparison of actual and synthetic estimated ISO polarizabilities at Pole Mountain 23
16 Example fit of a sum of exponential decays to estimated polarizabilities for an ISO target 24
17 MetalMapper data predicted using unsmoothed and smoothed polarizability estimates 25
18 Fits to noise distribution derived from ISO target polarizabilities, for channel 42 of MetalMapper data at Pole Mountain 25
19 Estimated normal mixture model parameters at each MetalMapper channel 26
20 Structure of the noise covariance matrices at Camp Beale and Pole Mountain 28
21 Effect of target depth on polarizability uncertainty 29
22 Dependence of polarizability standard deviations on horizontal target location 30
23 Predicted ISO polarizabilities for four spatial distribution scenarios 31
24 ISO polarizabilities for TEMTADS2x2 and BUDhh handheld sensors at Camp Beale 32
25 Standard deviation of positional error as a function of target position 34
26 Targets used for ROC prediction 37
27 ROC prediction for the MetalMapper, Pole Mountain noise 37
28 ROC prediction for the MetalMapper, Camp Beale noise. 38
29 Effect of target orientation on decision statistic and ROC 39
30 Numerical simulations of primary polarizability estimation for vertical and horizontal ISO targets. 40
31 Total polarizabilities estimated for ISO targets at Camp Beale. 42
32 EM-61 size-decay feature space, Camp Beale handheld area. 43

33 Dependence of EM61 total polarizability standard deviations at Camp Beale on cross track spacing 44
34 Dependence of lognormal decay distribution on target size and correlation coefficient 45
35 Performance prediction for the EM61, Camp Beale noise. 47

List of Tables

1 EM sensors used for UXO classification 7

List of Acronyms

BUD	Berkeley UXO Discriminator
BUDHH	Handheld Berkeley UXO Discriminator
CDF	Cumulative Density Function
cm	centimeter
DSS	Decision Support System
EM	Electromagnetic
EMI	Electromagnetic induction
ESTCP	Environmental Security Technology Certification Program
FAR	False Alarm Rate
FPF	False Positive Fraction
FLBGR	Former Lowry Bombing and Gunnery Range
GSV	Geophysical System Verification
ISO	Industry Standard Object
JDR	Jeep and Demolition Range
m	Meter
mm	Millimeter
MPV	Man-Portable Vector Sensor
MSEMS	Man-Portable Simultaneous EMI and Magnetometer System
ms	Millisececond
MTADS	Multi-Sensor Towed Array Detection System
mV	MilliVolt
PI	Principal Investigator
POC	Point of Contact
QC	Quality Control
ROC	Receiver Operating Characteristic
SLO	San Luis Obispo
SNR	Signal to Noise Ratio
TEM	Time-domain electromagnetic
TEMTADS	Time Domain Electromagnetic Towed Array Detection System
TOI	Target of Interest
TPF	True Positive Fraction
UBC-GIF	University of British Columbia Geophysical Inversion Facility
UXO	Unexploded Ordnance

1. Introduction

While significant advances have been made in the acquisition and processing of geophysical data for classification of buried munitions, the success of any classification strategy strongly depends upon site characteristics, including range of munitions types and clutter, geological background, topography and vegetation. In this project our goal is to develop and validate the components of a decision support system (DSS) that will help site-managers and teams design surveys and data processing strategies to achieve optimal classification performance at the lowest attainable cost for a given site.

2. Background

Advanced classification of buried munitions requires a number of steps:

(1) Data acquisition: detection of buried targets with a geophysical sensor.
(2) Feature extraction: characterization of each target with features estimated through inversion of a parameterized physics based forward model.
(3) Classification: prioritization of targets for digging using estimated features.

The ultimate goal of this processing is to identify all targets of interest with a minimal number of false alarms.

2.1. **Data acquisition.** In the data acquisition stage, a geophysical sensor is deployed at the site. Time-domain electromagnetic (TEM) sensors are most commonly used for detection of buried metallic targets. These instruments actively transmit a time-varying primary magnetic field which illuminates the earth. The variation of the primary field induces currents in buried targets and these currents in turn produce a secondary field which can be measured by a receiver at the surface (figure 1)

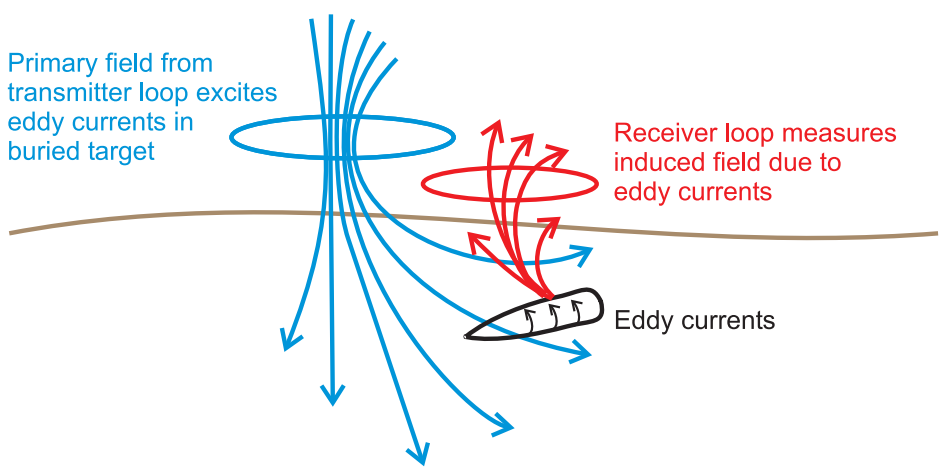


Figure 1. Electromagnetic induction (EMI) survey. Eddy currents are induced in a buried target by a time-varying primary field. Decaying secondary fields radiated by the target are then measured by a receiver at the surface.

In a detection mode survey, a TEM sensor passes over an area in nominally straight, parallel lines, with line spacing and instrument height dictated by instrument geometry and

detection considerations. Subsequent cued interrogations may revisit previously-identified targets and acquire finely spaced, high signal-to-noise ratio (SNR) data in a small area about the target. Recently developed systems for cued-interrogation illuminate the target with multiple transmitters and receivers (a "multistatic" configuration) from a single observation location and thereby avoid the requirement for accurate positioning of moving sensors. Table 1 compares the industry-standard EM-61 sensor with newer multistatic systems developed specifically for UXO detection and classification.

2.2. **Feature extraction.** Once target anomalies have been identified in the observed geophysical data, we can characterize each anomaly by estimating features which will subsequently allow a classification algorithm to discern targets of interest (TOI) from non-hazardous clutter (non-TOI). These features may be directly related to the observed data (e.g. anomaly amplitude at the first time channel), or they may be the parameters of a physical model. Advanced classification relies upon physics-based modeling, with the observed magnetic field $\mathbf{B}(t)$ radiated by a buried target usually represented as a time-varying dipole

$$\frac{\partial \mathbf{B_s}}{\partial t}(\mathbf{r},t) = \frac{p(t)}{r^3}\left(3(\hat{\mathbf{p}}(t)\cdot\hat{\mathbf{r}})\hat{\mathbf{r}} - \hat{\mathbf{p}}(t)\right) \tag{1}$$

with $\mathbf{r} = r\hat{\mathbf{r}}$ the separation between target and observation location, and $\mathbf{p}(t) = p(t)\hat{\mathbf{p}}(t)$ a time-varying dipole moment

$$\mathbf{p}(t) = \frac{1}{\mu_o}\mathbf{P}(t)\cdot\mathbf{B_o}. \tag{2}$$

The induced dipole is the projection of the primary field $\mathbf{B_o}$ onto the target's polarizability tensor $\mathbf{P}(t)$ (Bell et al., 2001). Here the elements of the polarizability tensor ($P_{ij}(t)$) represent the convolution of the target's B-field impulse response ($\tilde{\mathbf{P}}(t)$) with the transmitter waveform $i(t)$ (Wait, 1982)

$$P_{ij}(t) = \frac{\partial}{\partial t}\int_{-\infty}^{\infty} \tilde{P}_{ij}(t'-t)i(t')dt'. \tag{3}$$

The polarizability tensor is assumed to be symmetric and positive definite and so can be decomposed as

$$\mathbf{P}(t) = \mathbf{A}^T\mathbf{L}(t)\mathbf{A} \tag{4}$$

with $\mathbf{A}$ an orthogonal matrix which rotates the coordinate system from geographic coordinates to a local, body centered coordinate system. The diagonal eigenvalue matrix $\mathbf{L}(t)$ contains the principal polarizabilities $L_i(t)$ ($i = 1, 2, 3$), which are assumed to be independent of target orientation and location.

From a set of observations of the electromagnetic field, the inverse problem is then to find the set of model parameters (location, orientation, and polarizabilities) that best fits the data. The model vector $\mathbf{m}$ can be estimated by minimizing a norm (e.g. least squares) quantifying the misfit between observed ($\mathbf{d}^{obs}$) and predicted ($\mathbf{d}^{pred}$) data. For the TEM

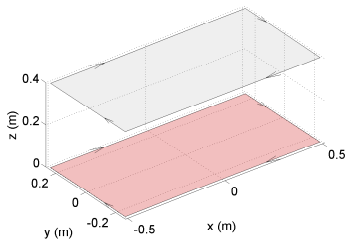
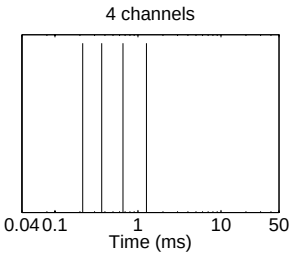
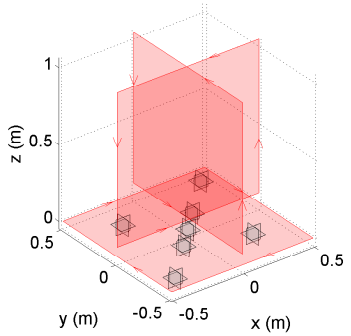
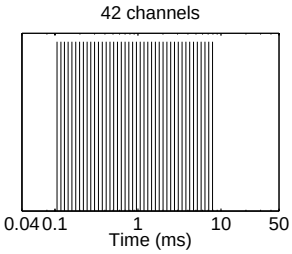
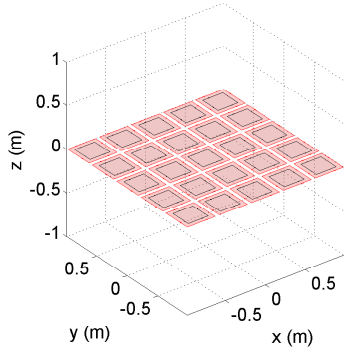
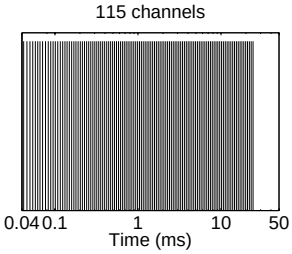
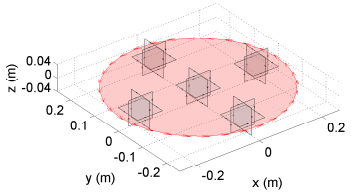
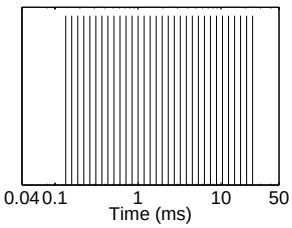
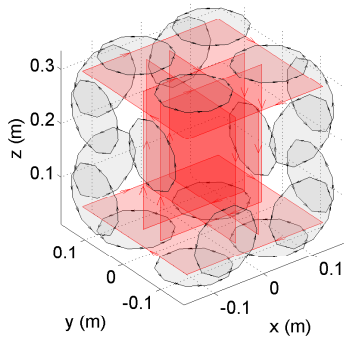
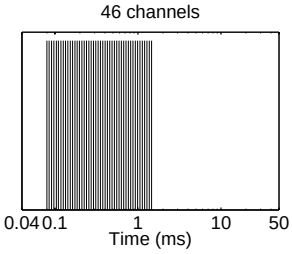
Sensor	Geometry	Channels
EM61 MKII		
MetalMapper		
TEMTADS		
MPV		
BUDHH		

TABLE 1. Electromagnetic sensors used for UXO classification. In geometry plots transmitters and receivers are red and black, respectively.

dipole model, the least squares estimate must generally be obtained iteratively, owing to the nonlinear relationship between model parameters and predicted data in equation 1. However, if the location of the target ($\mathbf{r}$) is assumed known, then the forward modeling becomes linear, so that

$$\mathbf{d}^{pred} = \mathbf{G}\mathbf{m} \tag{5}$$

with $\mathbf{G}$, the forward modeling matrix, implicitly dependent on target location. The least squares model estimate is then given by

$$\begin{aligned} \hat{\mathbf{m}} &= (\mathbf{G}^T\mathbf{G})^{-1}\mathbf{G}^T\mathbf{d}^{obs} \\ &= \mathbf{G}^{\dagger}\mathbf{d}^{obs} \end{aligned} \tag{6}$$

with

$$\mathbf{G}^{\dagger} = (\mathbf{G}^T\mathbf{G})^{-1}\mathbf{G}^T \tag{7}$$

denoting the pseudo-inverse. When inverting observed field data for a sensor with tri-axial transmit and receive coils (e.g. MetalMapper), we express $\mathbf{G}$ as

$$\mathbf{G}(\mathbf{r}) = \begin{bmatrix} B_s^x B_p^x \\ B_s^x B_p^y + B_s^y B_p^x \\ B_s^x B_p^z + B_s^z B_p^x \\ B_s^y B_p^y \\ B_s^y B_p^z + B_s^z B_p^y \\ B_s^z B_p^z \end{bmatrix}^T \tag{8}$$

with B_p the primary field at the target and B_s the secondary field at the receiver, with all fields implicitly dependent upon target ($\mathbf{r}$) and sensor location. Superscripts denote the x, y, z components of the respective fields. In this formulation, the model vector at location $\mathbf{r}$ is parameterized in terms of the six unique elements of the polarizability tensor $\mathbf{P}$

$$\mathbf{m} = [P_{xx},\ P_{xy},\ P_{xz},\ P_{yy},\ P_{yz},\ P_{zz}]^T\ . \tag{9}$$

In practice, the vector $\mathbf{m}$ is estimated at each time channel in a sequential, or two stage, inversion strategy (Song et al., 2011). As illustrated in figure 2, we first solve a nonlinear inverse problem for target location. We then solve a linear problem for the polarizability tensor elements (equation 9) at our fixed location estimate. Decoupling the location and polarizability parameters in this way allows for efficient parallel solution of the linear problem at all time channels. Target orientation and principal polarizabilities are subsequently estimated via joint diagonalization (Cardoso, 1996). This algorithm returns a single eigenvector matrix $\mathbf{A}$ for all channels, corresponding to a fixed target orientation. The eigenvalues at each time channel are then an estimate of principal polarizabilities.

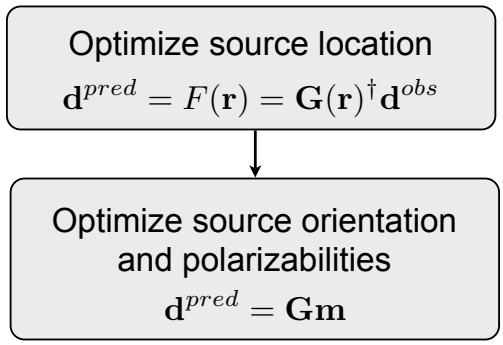


FIGURE 2. Sequential inversion approach for estimation of dipole model parameters. We first estimate target location $\mathbf{r}$; the predicted data in this case are a nonlinear functional $F(\mathbf{r})$. We then estimate target orientation and polarizabilities. At a fixed location and orientation, the predicted data are related to the model via a linear forward operator $\mathbf{G}$.

The model vector at each time channel can also be comprised of the principal polarizabilities

$$\mathbf{m} = [L_1,\ L_2,\ L_3]\,. \tag{10}$$

In this case the matrix $\mathbf{G}$ depends implicitly upon both target location and orientation. As will be discussed in section 4, this formulation of the model vector and forward operator is useful when propagating noise on the data to predicted distributions of polarizabilities.

2.3. **Classification.** To rank detected targets for digging, we use the information in our observed geophysical data. Features of the observed data, estimated without resorting to inversion with a physics-based model, can sometimes suffice as criteria to discriminate between ordnance and non-ordnance targets (e.g. Williams et al. (2007)). However, because dipole model parameters can be related to intrinsic properties such as target size and shape, features derived from the estimated parameters (such as the total polarizability in equation 3) are often more reliable for discriminating between TOI and non-TOI.

Classification with TEM data is often performed by comparing estimated polarizabilities with library responses and then ranking a target based on some measure of closeness between observed and expected responses. Care must be taken here to use parameters which can be reliably estimated: late time polarizabilities are more susceptible to noise and poor estimates may unduly affect the discrimination decision. Pasion et al. (2007) solve this problem with a fingerprinting algorithm that inverts for target location and orientation while holding principal polarizabilities fixed at their library values. Reducing the model's degrees of freedom in this way makes the inversion less susceptible to fitting the noise. Targets are then dug based upon the proposed library item which produces the best fit to the observed data.

Statistical classification algorithms applied to UXO discrimination try to learn a decision rule from a sample of labeled targets for which ground truth is known (the training data).

One approach to formulating the decision rule is to fit some assumed parametric distributions to each class of targets in the training data, and to then assign an unlabeled (test) target to the class distribution which is most likely. The class distributions are defined in a multidimensional feature space spanned by some subset of estimated model parameters, or transformations thereof. The total polarizability

$$L_{total}(t_j) = \sum_{i=1}^{3} L_i(t_j) \tag{11}$$

can be a useful classification feature when secondary and tertiary ($L_2,\ L_3$) polarizabilities are poorly constrained. Alternatively, a two-dimensional space spanned by the amplitude and decay of the polarizabilities

$$\begin{aligned} \text{amplitude} &= \log_{10}\left(\sum_{j=1}^{N} L_{total}(t_j)\right) \\ \text{decay}(t_k, t_j) &= \frac{L_{total}(t_k)}{L_{total}(t_j)}, \end{aligned} \tag{12}$$

has been successfully used to train statistical classifiers for a number of ESTCP field demonstrations (e.g. Billings et al. (2010)). These parameters are useful because, to first order, a conductor can be modeled as a simple LR (circuit with inductor and resistor in series) loop which is inductively coupled to transmitters and receivers on the surface. The current response of this loop is a decaying exponential which is fully described by an amplitude and time constant (West and Macnae, 1991). Polarizabilities estimated from multistatic data are typically well-constrained relative to monostatic detection data, and so statistical classifiers for cued data sets are usually trained on the (log-transformed) principal polarizabilities, or total polarizabilities.

The output of any automated classification algorithm is a decision statistic, or score, that is used to rank detected targets from likely TOI to likely non-TOI. For example, a library classifier uses the misfit of estimated polarizabilities with library polarizabilities as the decision statistic. As shown in Figure 3, the receiver operating characteristic (ROC) is then a plot of the true positive fraction (TPF) versus the false positive fraction (FPF), which are defined as the cumulative score distributions for TOI and non-TOI

$$\begin{aligned} TPF(x) &= \int^{x} p(y|TOI)dy \\ FPF(x) &= \int^{x} p(y|non-TOI)dy. \end{aligned} \tag{13}$$

In the context of munitions management, the false alarm rate (FAR) at which all ordnance are detected on the ROC (i.e. $TPF = 1$) is the crucial metric by which site managers evaluate the efficacy of remediation efforts. An advanced technique that results in good initial detection of TOI but fails to find outlying TOI until late in the dig order will be

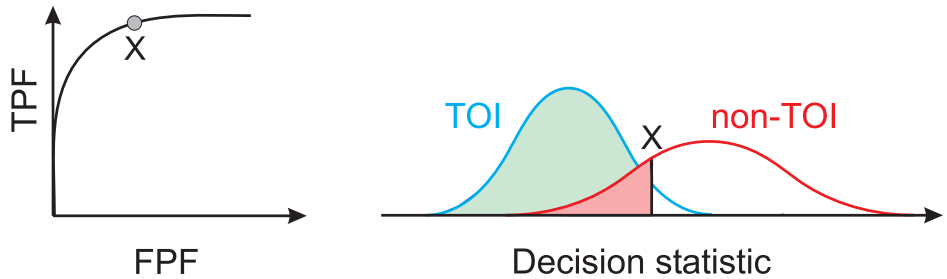


FIGURE 3. Left: the receiver operating characteristic (ROC) shows the true positive fraction (TPF) as a function of the false positive fraction (FPF). Right: the ROC at point x can be modeled as the integral of the distributions of TOI and non-TOI (true and false positives, respectively) with respect to the decision statistic.

judged unsuccessful. This can occur with statistical classifiers that overfit the training data and so fail to generalize to noisy test TOI.

3. Predicting target response thresholds

Targets identified in a detection mode survey must be selected for subsequent cued interrogation using objective and reproducible criteria. The Geophysical System Verification (GSV) approach developed by ESTCP (Nelson et al., 2010) advocates setting a target picking threshold based upon the lowest predicted amplitude expected for targets of interest at a specified clearance depth. In the case of the EM-61 cart, the worst case scenario is a horizontal item, oriented perpendicular to the sensor track. This geometry decouples the primary field from the axial (L_1) response of the target, so that the dipole response is dominated by the sum of the smaller, transverse polarizabilities. The GSV response calculator provides functionality for predicting the best and worst case target response amplitudes for an array of EM-61 sensors in an arbitrary geometry.

When we apply this approach to munitions response projects, we sometimes find that the threshold defined by the worst case scenario results in an undesirably large number of target picks. This has motivated us to extend the GSV response calculator functionality to consider the statistics of the target response across the swath of the array.

Figure 4 illustrates this extended analysis for a three coil EM-61 array interrogating a 37 mm projectile at 30 cm depth. Within the swath of the array, the minimum predicted anomaly amplitude occurs for orientations of the target across the sensor track. The global minimum of the response ($\approx$ 8.5 mV at channel 3) is near the cross-track edges of the center transmitter coil. This minimum is the consequence of the fall-off in received EMF as the secondary dipole moves away from the center of each receiver coil. In this particular example, the minimum response is in a cross track orientation.

For asymmetric array geometries we have found that the minimum response is still at a horizontal dip, but the target azimuth can be at an intermediate angle to cross and along track directions. This is illustrated in figure 5 for an asymmetric three coil array. When constructing the worst case curve we therefore model over a full 360 degree range

of target azimuths to obtain the minimum predicted response. Comparing the symmetric and asymmetric array geometries in figures 4 and 5, we remark that the minimum response across the swath is significantly elevated for the asymmetric array. Narrowing the swath will decrease the production rate, but will improve target detection.

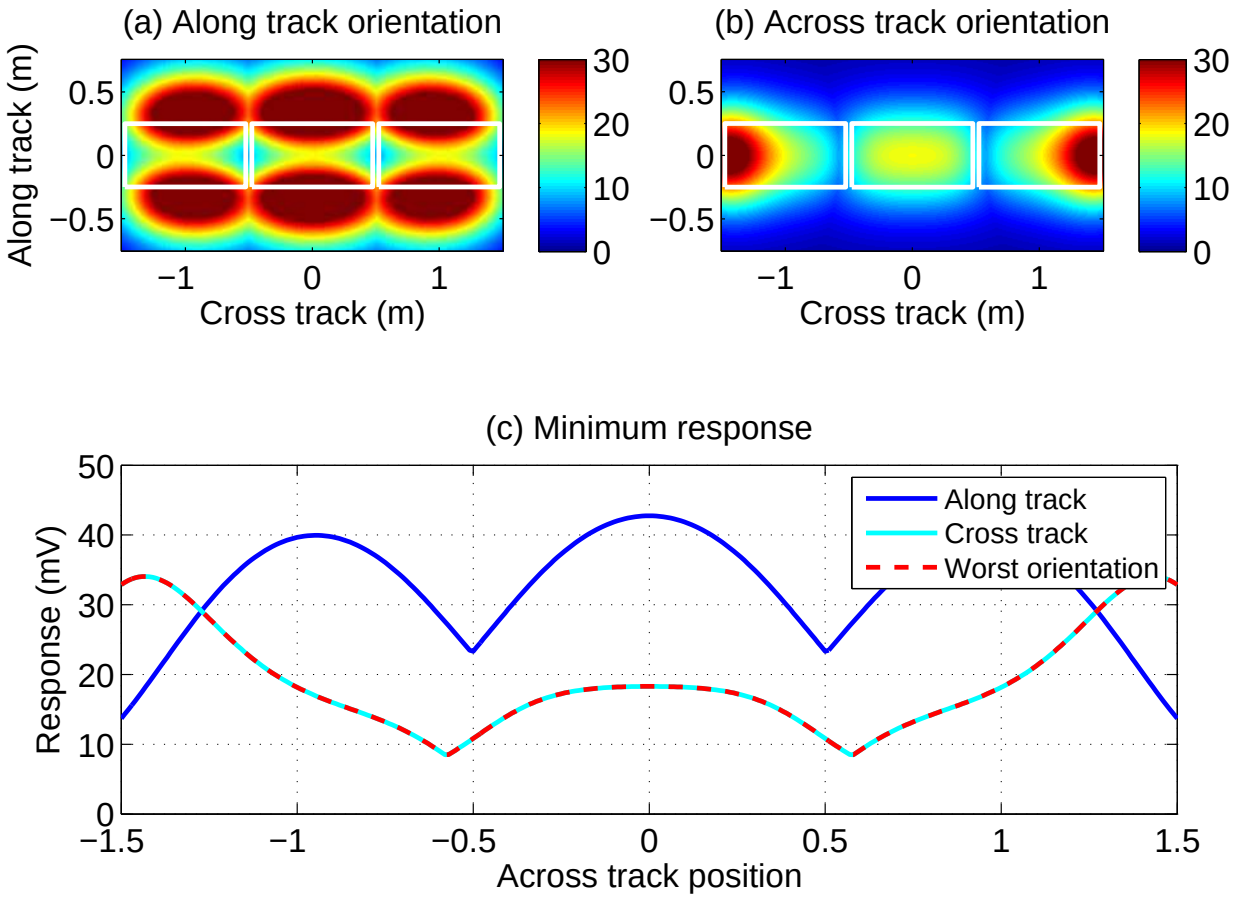


FIGURE 4. Anomaly amplitude analysis for a horizontal 37 mm projectile at 30 cm depth illuminated by a three coil EM-61 array. (a) Maximum anomaly amplitude (mV) at channel 3 as a function of target location, for target oriented along track. White lines indicate transmitters. (b) As in (a), but with horizontal target oriented across track. (c) Minimum of responses in (a) and (b) across the swath of the sensor array.

Returning now to the symmetric three coil array, we note that the 8.5 mV minimum in figure 4(c) will undoubtedly produce a very large number of picks that must be excavated or interrogated with a cued sensor. For example, at the Jeep and Demolition Range (JDR) at the Former Lowry Bombing and Gunnery Range (FLBGR), we found that a 30 cm clearance depth for 37 mm projectiles, corresponding to an 8.5 mV threshold, required us to dig a prohibitively large number of picks.

However, we have found that by considering the statistics of the target response we can significantly reduce the number of target picks. Assuming a uniform spatial distribution of horizontal 37 mm projectiles throughout the sensor footprint with a uniform distribution of target azimuths, we obtain the cumulative distribution of response amplitudes in figure 6.

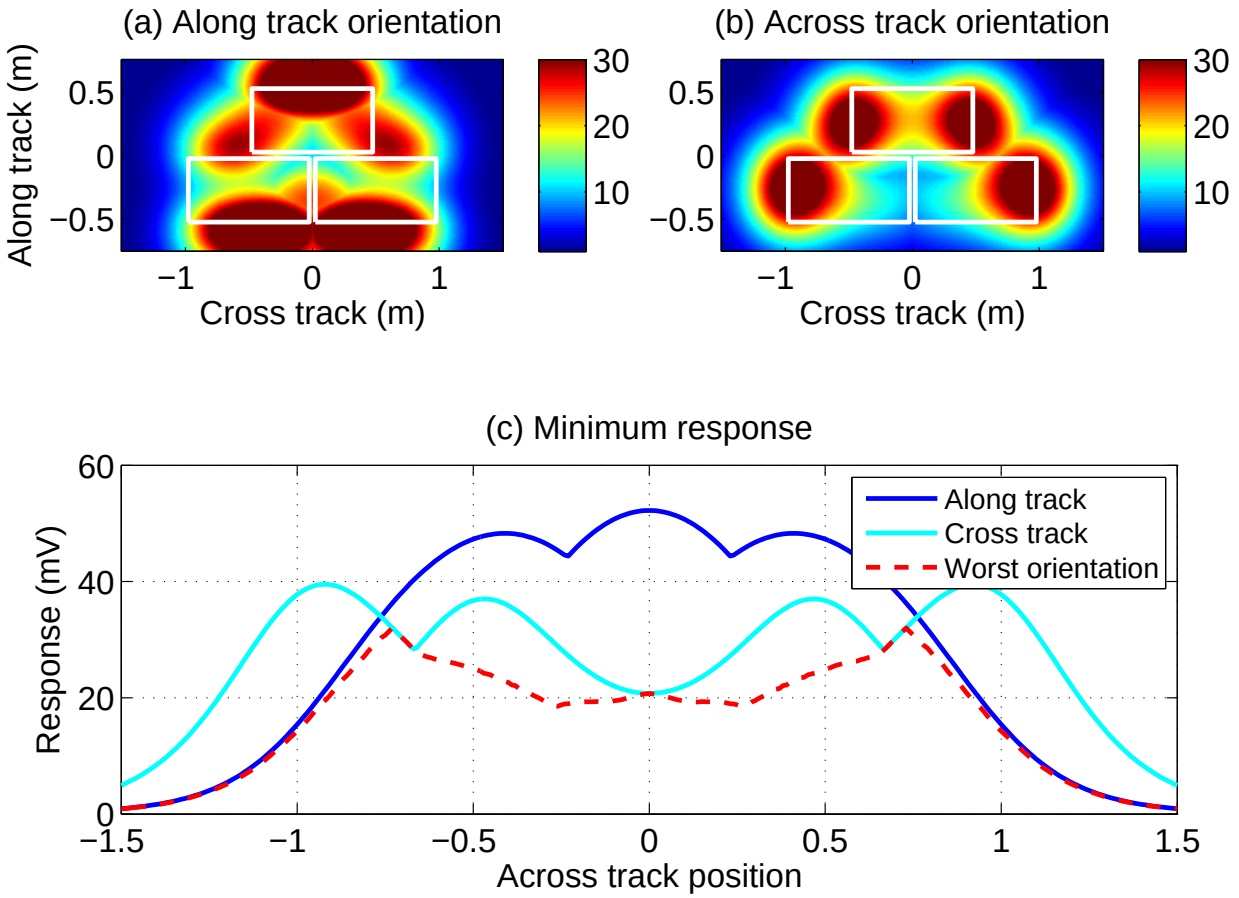


FIGURE 5. Anomaly amplitude analysis for an asymmetric three coil EM-61 array. Note in this case that the worst case amplitude does not always correspond to the cross track orientation.

At 0.99 and 0.95 confidence levels, we can raise the detection threshold in figure 4 to approximately 11 and 15 mV, respectively. In (c) we plot the proportion of 37 mm targets that fall below these thresholds as a function of cross track location. For the 11 mV threshold (0.99 confidence level), the non-zero regions of targets below the threshold are restricted to lobes near the outer edges of the center transmitter. Increasing the threshold to 15 mV (0.95 confidence level) significantly increases the chances that we will miss a 37 mm target within the sensor swath.

As shown in figure 7, increasing the threshold to 11 mV at JDR reduced the number of target picks by 30%, while still ensuring a very high confidence of finding all 37 mm projectiles. At the 95% confidence threshold we have a 60% reduction in target picks. However, the 5% probability of missing a horizontal 37 mm may not be an acceptable risk. At sites where the worst case threshold produces a prohibitive number of picks, this type of statistical analysis can provide objective justification for slightly raising the detection threshold.

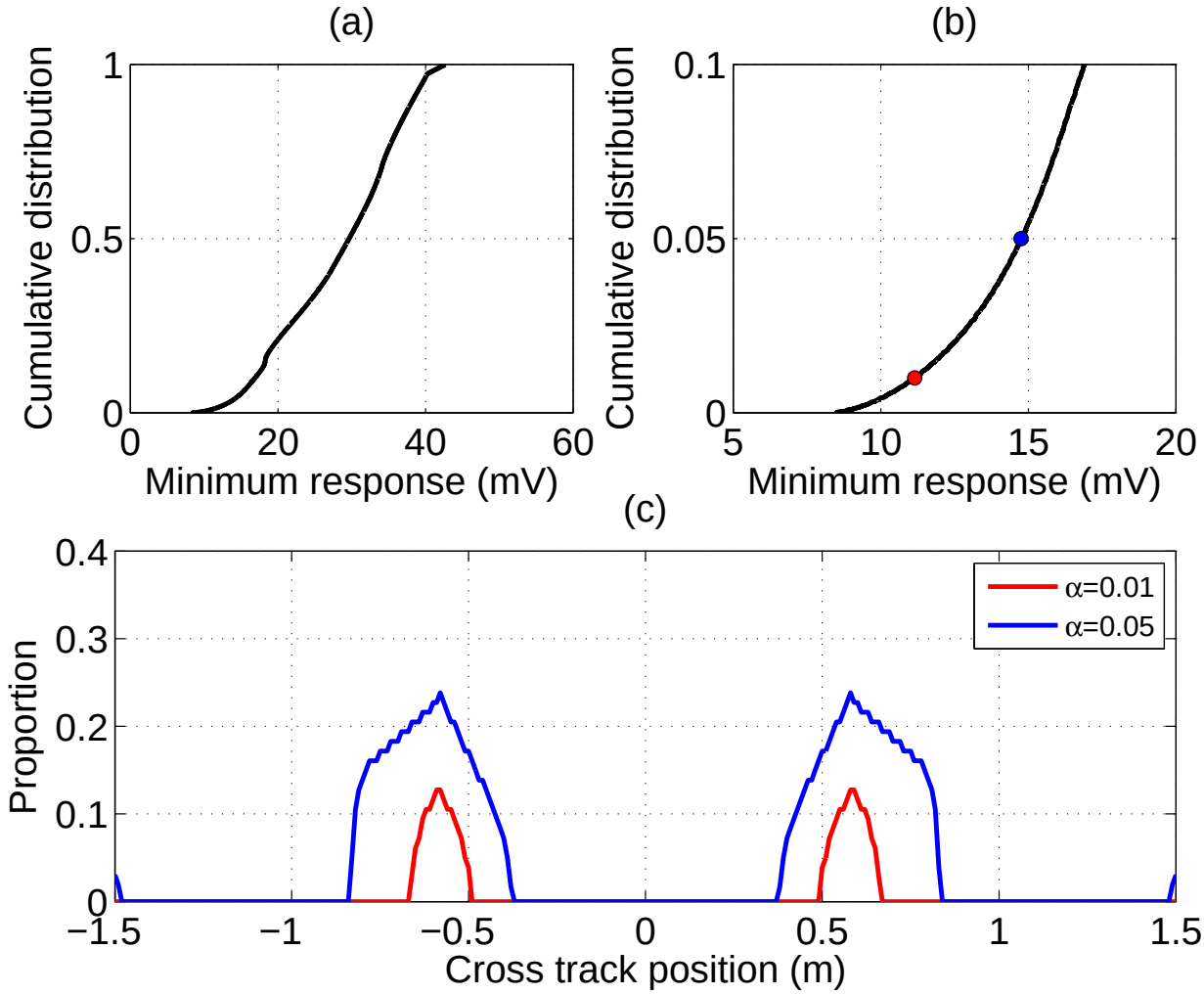


FIGURE 6. (a) Cumulative distribution of response amplitudes for horizontal 37 mm projectiles at 30 cm depth. Targets are distributed uniformly with respect to cross track location and azimuth. (b) Lower tail of the cumulative distribution. Markers indicate 0.99 and 0.95 confidence bounds. (c) Proportion of horizontal 37 mm targets with response below threshold specified by confidence level $1 - \alpha$, as a function of cross track location.

The distribution in figure 6 is generated for horizontal 37 mm targets at arbitrary azimuths. Allowing for an arbitrary target azimuth and dip generates a somewhat less conservative distribution of response amplitudes. Figure 8 shows the empirical probability and cumulative density functions for simulations assuming a uniform angular distribution of targets. These results are again generated for 37 mm projectiles at 30 cm depth, with the symmetric EM-61 array of figure 4.

In future work we will extended this type of GSV functionality to newer dynamic platforms, including the MetalMapper and TEMTADS2x2. If requested by the program office, we will work to develop a software tool, similar to the existing GSV calculator, that makes this analysis widely available to the munitions response industry.

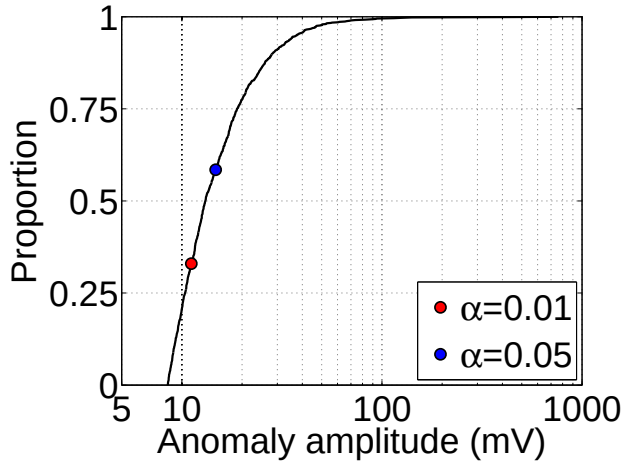


FIGURE 7. Cumulative distribution of response amplitudes at JDR. Markers indicate $1-\alpha$ confidence thresholds for detection of 37 mm targets.

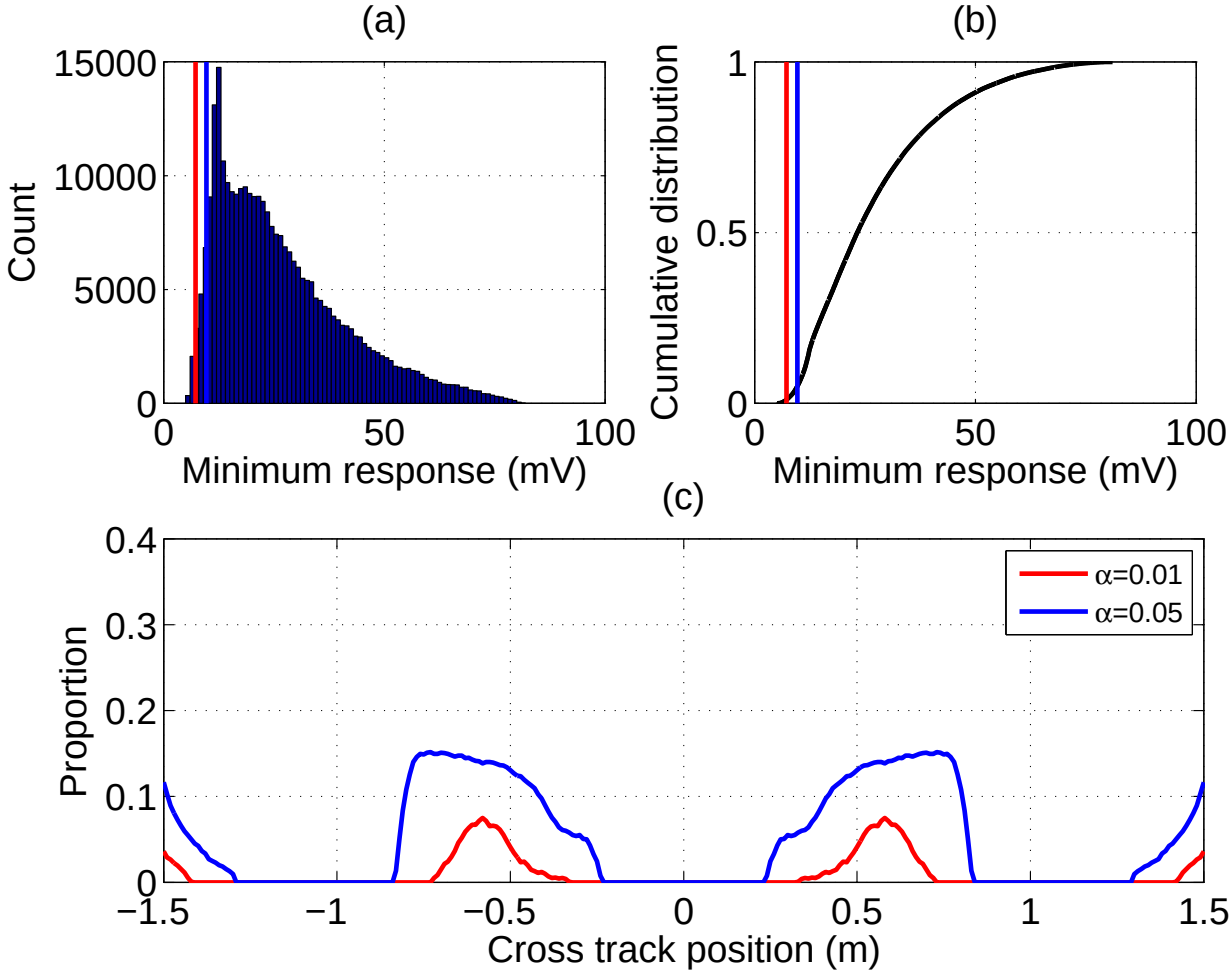


FIGURE 8. Distribution of response amplitudes for 37 mm targets at 30 cm depth, distributed uniformly in azimuth and dip. (a) Empirical distribution. Vertical lines are $1-\alpha$ confidence level cut-off. (b) Cumulative distribution. (c) Proportion of 37 mm targets with response below threshold specified by $1-\alpha$ confidence level, as a function of cross track location.

4. Performance prediction

Excellent results have been obtained throughout the ESTCP demonstration program with the sensor and data processing technologies described above: advanced classification consistently outperforms conventional processing of industry-standard detection data (Billings et al. (2010), Steinhurst et al. (2010), Prouty et al. (2011), Shubitidze et al. (2011)). The significant reductions in the false alarm rate obtained with advanced classification translate to substantial cost-savings during site remediation. Motivated by these successes, classification methods have increasingly been transitioned from researchers to industry. For example, at the 2011 Pole Mountain demonstration, production geophysicists working with UBC-GIF software to process MetalMapper data obtained classification performance comparable to that of "expert" analysts (Pasion, 2012). These initial results indicate that advanced classification can be fully transitioned to the munitions response industry.

To further this transition process, we have developed methods for efficient prediction of classification performance. The goal of this work is to provide tools for site managers and their teams to reliably and objectively predict the benefits of advanced classification for a particular remediation problem. In particular, we address two questions:

(1) Will multi-static sensor data reliably discriminate between targets of interest and representative clutter items encountered at the site?
(2) Is there a significant improvement in expected classification performance - and hence a reduction in expected costs - relative to classification with a monostatic sensor such as the EM-61? For a relatively simple classification task (e.g. identification of 4.2" mortars at Camp Sibert, Billings et al. (2010)), the additional survey and processing costs associated with advanced classification may ultimately provide negligible improvements in classification.

Throughout this report we will exploit available data sets from ESTCP demonstrations to develop and validate our approach to performance prediction. A summary of these data sets is provided in Appendix A.

As discussed in the previous section, survey design for munitions response projects is largely restricted to detection considerations. Here we shift the emphasis in survey design from detection to classification. We consider how site-specific inputs (sensor platform, target type, etc.) ultimately affect classification performance, as quantified by the ROC. Two routes can be taken to map from these inputs to a predicted ROC curve. First, numerical simulation of synthetic observed sensor data can be used to quantify the variability of estimated polarizabilities, and hence to predict the score distributions of TOI and non-TOI for a given noise model. This approach is numerically intensive and is not suited for an efficient decision support tool. We therefore pursue the second option: we develop statistical mappings between environmental inputs and estimated polarizabilities, as depicted in figure 9.

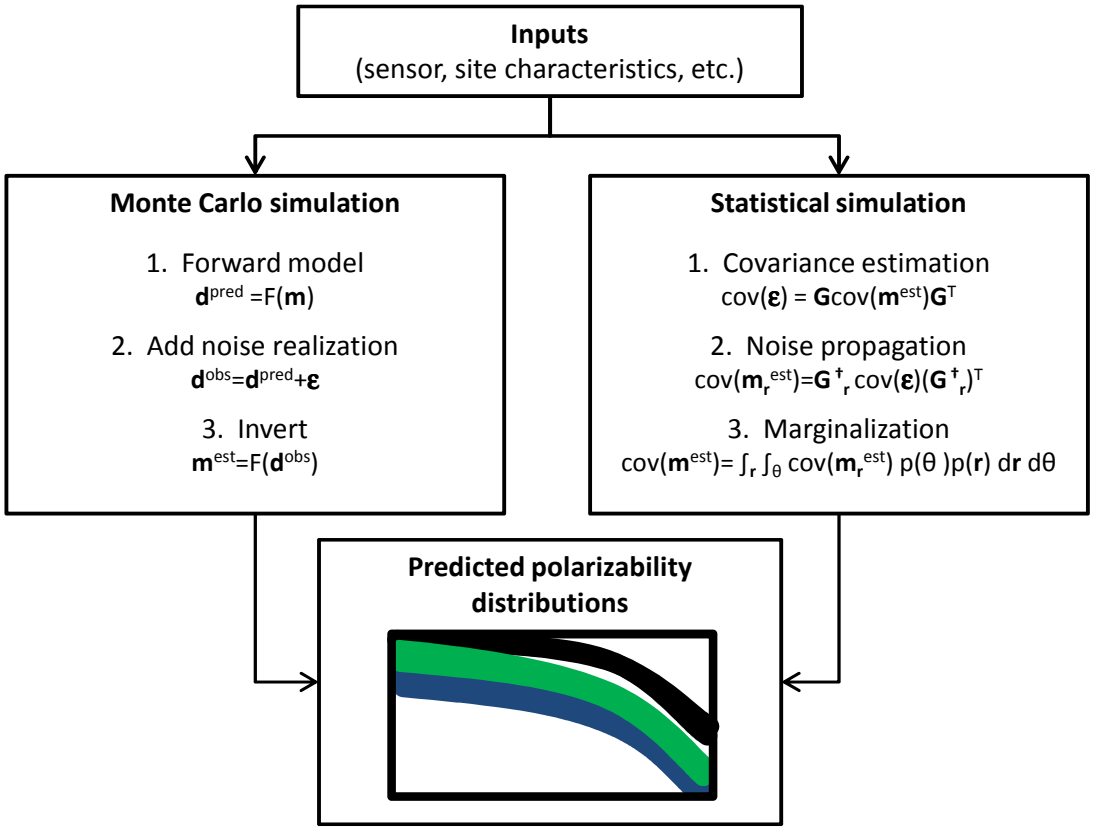


FIGURE 9. Methods for polarizability prediction. Most studies use inversion of synthetic data with independent, Gaussian noise to predict polarizability distributions. In the next section, we develop statistical mappings that characterize correlated, non-Gaussian noise with a noise covariance, then propagate noise to the model and integrate over specified spatial and angular distributions of targets.

4.1. **Performance prediction for cued interrogation.**

4.1.1. *Noise estimation.* We begin by quantifying the noise on the observed data for a stationary cued sensor such as the MetalMapper. The problem is somewhat simplified by the fact that soundings are acquired at a single location; we need not account for the errors in measurement of sensor location that complicate analysis of dynamically-collected data. The errors on measured cued data are then a combination of sources, including baseline sensor noise and EMI responses that are unaccounted for in an inversion (e.g. transient cultural noise, magnetic soils or neighboring targets).

Background noise is characterized via repeated measurements taken at an electromagnetically quiet location at the site. These measurements quantify sensor noise and soil response. Significant variability in sensor noise was observed with two MetalMapper systems used for the 2010 Camp Butner demonstration. The older, prototype system had substantially higher noise levels than the newer system manufactured by Geometrics. This difference was likely due to improvements in electronics in the newer system. Subsequent deployments of Geometrics MetalMappers at Pole Mountain and Camp Beale have produced consistently

low noise levels associated with the sensor platform and electronics. Figure 10 compares standard deviations estimated from background measurements taken at Camp Butner and Pole Mountain.

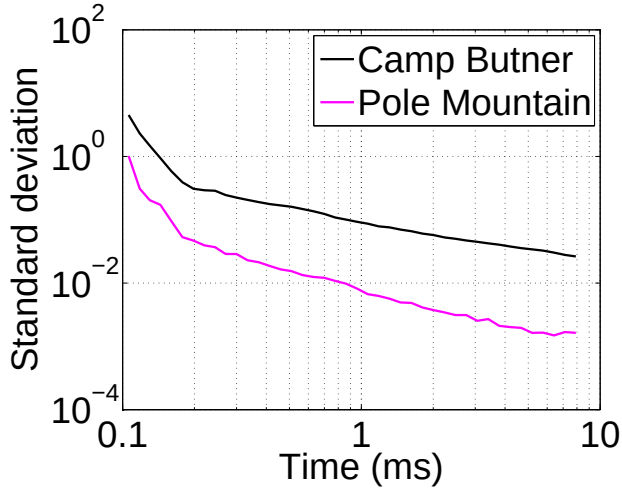


FIGURE 10. Background noise standard deviations for MetalMapper sensor at Camp Butner and Pole Mountain.

When inverting field data, we assume that the observed data are a superposition of the true data predicted by one or more dipole sources, and a noise term ϵ

$$\mathbf{d}^{obs} = \mathbf{d}^{true} + \epsilon \tag{14}$$

with $\epsilon \sim \mathcal{N}(0, \sigma_i)$ a zero mean, normally-distributed random variable. The estimated standard deviations weight the contributions of each channel to the data misfit

$$\phi_d = \sum_i \left(\frac{d_i^{obs} - d_i^{true}}{\sigma_i} \right)^2 . \tag{15}$$

In the sequential inversion approach depicted in figure 2, this weighting is less critical since each channel is inverted separately at a fixed location. Channel weighting is more important when all channels are inverted at once, as in the fingerprinting approach described in Pasion (2007).

Of course, the estimated standard deviations derived from background measurements do not account for the various non-stationary sources of noise that are inevitably present in field data. Figure 11 shows one such source: an elevated magnetic soil response encountered in part of the survey area at Camp Butner. Non-dipolar fields that cannot be fit in an inversion can also be regarded as additive noise on the observed data. Given the complex dependence of noise upon sensor and environment, we do not attempt to predict individual noise sources and their subsequent effect on estimated polarizabilities. Rather, we use the distributions of polarizabilities recovered from ESTCP field data sets to estimate the *effective* noise across each site.

Figure 12 compares polarizabilities recovered for small industry standard objects (ISOs) at Pole Mountain and Camp Beale. It is evident from the distributions of polarizabilities

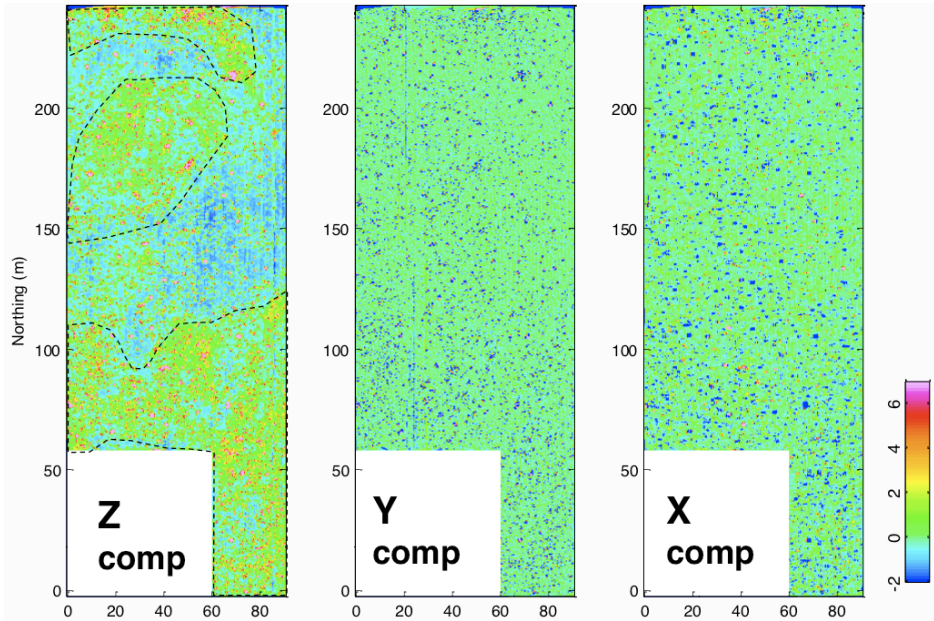


FIGURE 11. Components of the observed secondary field for MetalMapper dynamic data acquired at Camp Butner. Regions of elevated magnetic soil response are delineated in the z-component receiver data. Horizontal components (x and y) of the data are less sensitive to magnetic soils when excitation is from a z-component transmitter only.

that the effective noise at Camp Beale is larger than at Pole Mountain. Assuming perfect recovery of target location, each set of estimated polarizabilities in Figure 12 is a linear transformation of the observed data acquired over an ISO target (see equation 6). The covariance of the jth polarizability model at the ith time channel can then be expressed as

$$\mathrm{cov}(\mathbf{m}_{ij}) = (\mathbf{G}_j^{\dagger})^T \mathrm{cov}(\mathbf{d}_i^{obs}) \mathbf{G}_j^{\dagger} \tag{16}$$

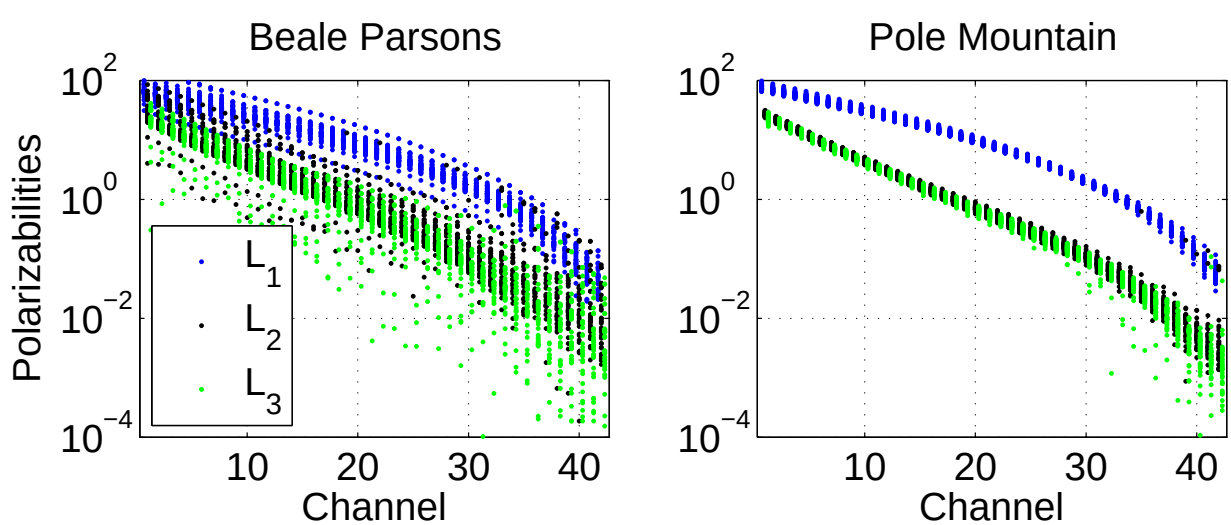


FIGURE 12. Estimated ISO polarizabilities at Camp Beale and Pole Mountain.

with the pseudo-inverse $\mathbf{G}_j^\dagger$ given by equation 7. For uncorrelated errors with constant standard deviation σ_i at each time channel, the data covariance is $\text{cov}(\mathbf{d}_i^{obs}) = \sigma_i^2\mathbf{I}$ and the model covariance simplifies to

$$\text{cov}(\mathbf{m}_{ij}) = \sigma_i^2(\mathbf{G}_j^T\mathbf{G}_j)^{-1} \tag{17}$$

As discussed in section 2.2, the forward modeling operator $\mathbf{G}$ can be computed for a fixed target location and orientation, in which case the model vector at each time channel is composed of the principal polarizabilities at that channel (equation 10).

The covariance of the principal polarizabilities in equation 17 represents the covariance of the model parameters that would be estimated from repeated noise realizations with standard deviation σ_i for a single target at a fixed location and orientation. The polarizabilities estimated from field data are instead a set of model estimates with independent realizations of noise but at varying locations and orientations. The following procedure computes an estimator of the effective noise standard deviation at a site from the set of recovered principal polarizabilities:

Algorithm 1 Estimator $\hat{\sigma}_1^2$ of the data variance.

(1) Estimate the (3×3) sample covariance matrix using the recovered principal polarizabilities at i^{th} time channel, for all targets within a given class.
(2) Form a vector δ_i comprised of the diagonal elements of the estimated model covariance (i.e. the variances of the polarizabilities)
(3) At each target's estimated location and orientation, compute $\mathbf{\Gamma}_j = (\mathbf{G}_j^T\mathbf{G}_j)^{-1}$. This is the model covariance for the jth target (equation 17) without scaling by the (unknown) data variance σ_i^2
(4) Compute the mean model covariance matrix

$$\bar{\mathbf{\Gamma}} = \frac{1}{M}\sum_{j=1}^{M}\mathbf{\Gamma}_j \tag{18}$$

over all M targets in the class. $\bar{\mathbf{\Gamma}}$ is an average of symmetric and positive-definite covariance matrices and so is itself a valid covariance matrix.
(5) Form the vector $\mathbf{g}$ from the diagonal elements of $\bar{\mathbf{\Gamma}}$.
(6) Finally, compute a least squares estimate of the data variance at the ith time channel

$$\hat{\sigma}_i^2 = (\mathbf{g}^T\mathbf{g})^{-1}\mathbf{g}^T\delta_i \tag{19}$$

To ensure that the inverse problem in equation 19 is well-posed, we assume a diagonal data covariance and neglect estimation of correlations quantified by off-diagonal terms.

Figure 13 shows a synthetic example of noise estimation for MetalMapper data. The synthetic observed data are contaminated with additive zero mean Gaussian noise with standard deviations derived from Camp Butner background measurements. We simulate ISOs distributed uniformly in both location and orientation, with a maximum target depth of 0.6 m. The recovered distribution of polarizabilities, estimated from nonlinear inversion of the

synthetic data, is then used to compute the data variance estimator using the above procedure. There is good agreement between the actual and estimated data standard deviations, suggesting that this approach can provide a reasonable estimate of the effective noise on the data.

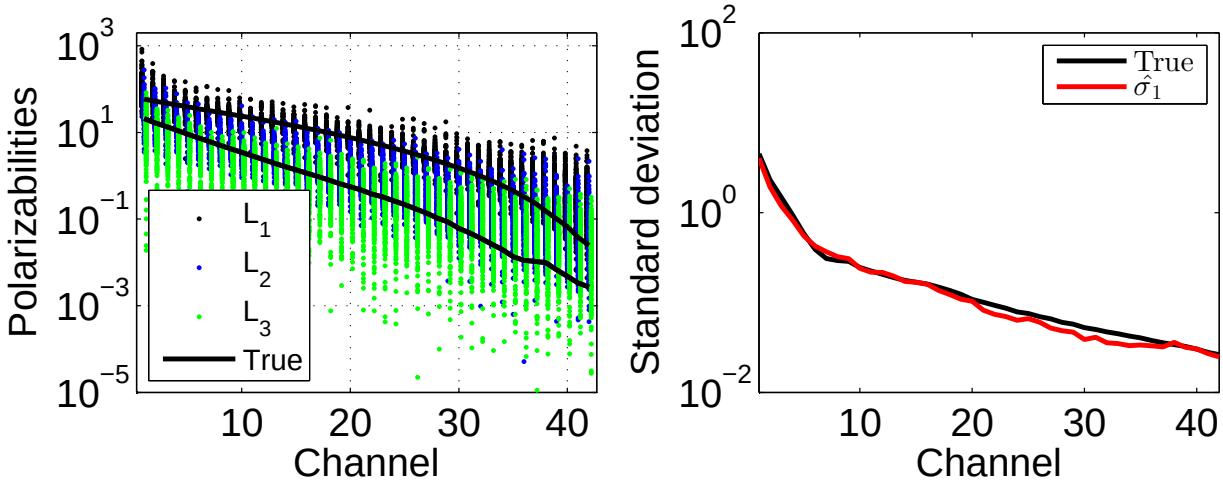


FIGURE 13. Synthetic example of estimation of effective noise standard deviation. Left: ISO polarizabilities estimated from synthetic MetalMapper data contaminated with Gaussian noise. The kink in the true secondary polarizabilities at late times is a consequence of noise on the particular measurements used to extract reference polarizabilities. Smoothing could be applied to remove this effect. Right: True standard deviation of noise and estimator $\hat{\sigma}_1$ recovered from polarizability distributions.

We now apply this analysis to real MetalMapper data. The analysis is restricted to ESTCP demonstration sites with seeded ISOs (at the time of this report): Pole Mountain and Camp Beale. ISO items are preferred for this analysis because of their physical uniformity and consequent consistency of their polarizabilities under low noise conditions. ISOs can also be seeded uniformly across the site to ensure that variability in site conditions is adequately sampled.

Figure 14 shows noise estimates for Camp Beale and Pole Mountain derived from the distributions of estimated ISO polarizabilities in figure 12. As expected from the observed distributions of polarizabilities, the estimated standard deviation of the noise at Camp Beale is larger than at Pole Mountain, but only slightly. Munkholm and Auken (1996) showed that log-gated and stacked white noise produces errors on the TDEM response with a standard deviation exhibiting a $t^{-1/2}$ decay. For these data sets, the estimated noise standard deviations decay considerably faster than t^{-1}.

As a check on this procedure, we can now contaminate synthetic data with our effective noise and verify that the distributions of recovered polarizabilities are consistent with the actual distributions in figure 12. Figure 15 compares the synthetic estimates recovered with the inferred Pole Mountain noise model and the actual estimated polarizabilities for ISOs at this site. The synthetic polarizabilities have a much higher variance, in particular for the

secondary and tertiary polarizability estimates, indicating that we have overestimated the noise. A similar result is obtained for simulations with our estimated Camp Beale noise.

Why does our noise estimation procedure fail in this case? We have made several assumptions in this analysis: the noise is independent, normally distributed, and constant for all data at each time channel. As we will now show, a closer examination of the polarizabilities estimated from field data indicates that the noise on the data is correlated, only approzimately normally-distributed, and is not constant for all data measured at a single time channel.

The correlated nature of the noise is easy to see in figure 12: correlated shifts of the estimated polarizabilities, especially obvious at Camp Beale, could not result from independent random noise. In the case of Camp Beale, a variable magnetic soil response affects our estimation of target depth and so results in constant shifts (in log space) of the estimated polarizabilities.

We turn now to the distribution of the noise. Typically, in synthetic simulation work we assume Gaussian noise at each time channel. This is motivated by simplicity, and the observation that multiple sources of additive noise will tend to a normal distribution.

In previous work, we also carried out simulations that treated the data residuals - the difference between observed and predicted data - as realizations of noise (Carin et al., 2012). The various realizations of noise across a site were then added to forward modeled data for a given target to synthetically seed that target at the site. This allowed us to study the variability of estimated target features expected at the site.

A drawback of the synthetic seeding method is that the residuals in an inversion in fact represent the component of the noise that is not fit in an inversion. We are not accounting for the component of the noise that causes random deviations of the estimated model and contributes to the overall variability of the features within a target class. Here we try to

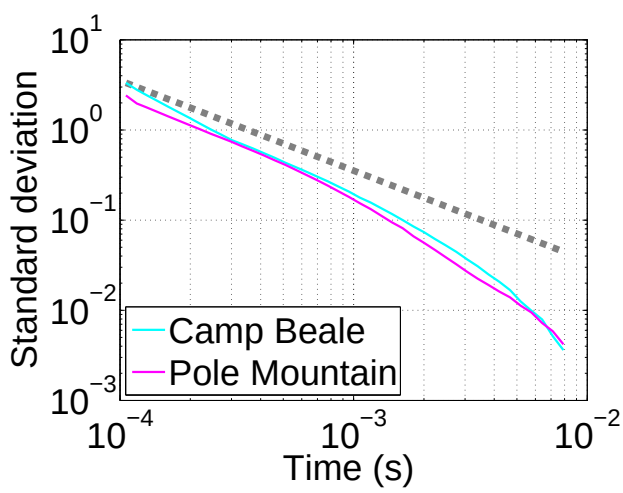


FIGURE 14. Estimates of effective noise standard deviations for MetalMapper Beale Parsons and Pole Mountain data sets. Heavy dashed line is a $t^{-1/2}$ decay.

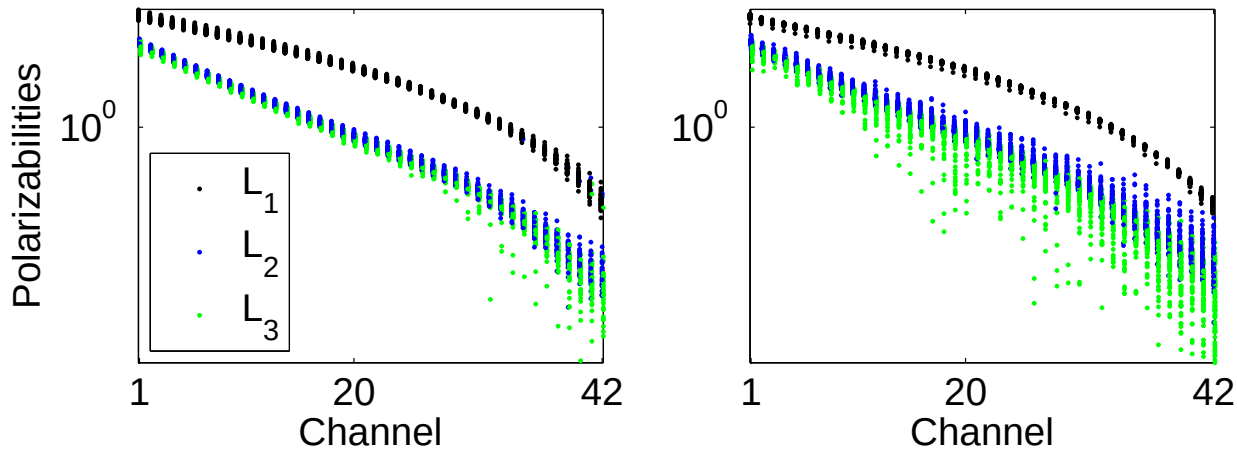


FIGURE 15. Comparison of actual (left) and synthetic estimated (right) ISO polarizabilities at Pole Mountain. The latter are generated by adding the predicted data and independent Gaussian noise with standard deviations estimated with algorithm 1.

isolate this effect of the noise by fitting a smooth, physically-motivated function to the estimated polarizabilities. This provides a route for estimating the component of the noise that contaminates our model estimates.

The polarizability decay for an arbitrary target can be generally parameterized as a superposition of decaying exponentials

$$L^{smooth}(t_j) = \sum_{i=1}^{N} a_i \exp(-t_j/\gamma_i). \tag{20}$$

With $N \to \infty$, this is the analytic form of the polarizability response of a spheroidal target (Kaufman, 1994). Random fluctuations of recovered polarizabilities away from this smoothly decaying function are then diagnostic of random noise on the data. We fit the above function to the recovered polarizabilities, with the summation truncated to $N = 30$ terms and decay constants γ_i logarithmically-spaced between 10^{-6} and 10^{-2} s. We solve a linear optimization problem for the coefficients a_i by minimizing the least squares difference between estimated and smoothed polarizabilites, with the constraint $a_i \geq 0$ ensuring a monotonically decreasing function. Figure 16 shows a typical fit obtained with this approach. There is excellent agreement between smoothed and unsmoothed polarizabilities, the only obvious deviation is at late times when the estimated secondary polarizabilities approach the noise floor. Given estimated (unsmoothed) and smoothed polarizabilities, we can generate an estimate of the noise by predicting the data for both

$$\begin{aligned} \mathbf{d}_{est} &= \mathbf{G}\mathbf{m}_{est} \\ \mathbf{d}_{smooth} &= \mathbf{G}\mathbf{m}_{smooth} \end{aligned} \tag{21}$$

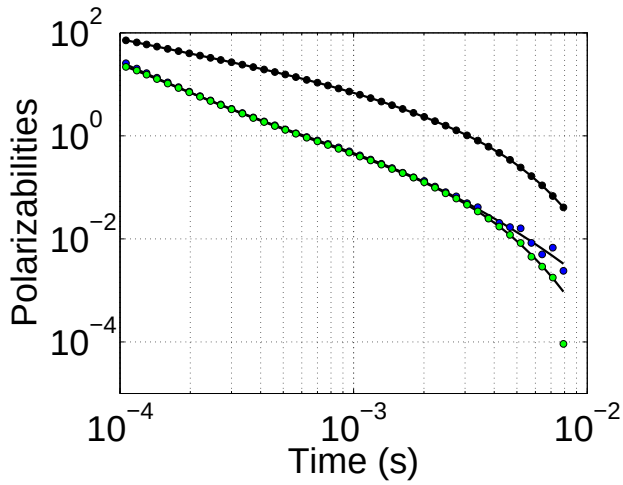


FIGURE 16. Example fit (solid lines) of a sum of exponential decays to estimated polarizabilities (markers) for an ISO target at Pole Mountain

where the model vectors $\mathbf{m}$ are the polarizabilities at each channel. Our prediction for the realization of noise on the data is then

$$\epsilon = \mathbf{d}_{est} - \mathbf{d}_{smooth} = \mathbf{G}\Delta\mathbf{m}. \tag{22}$$

From this we see that the predicted noise is that of a dipole source, co-located with the target and with principal polarizabilities given by the change in the polarizability model $\Delta\mathbf{m}$. Figure 17 shows the prediction for ϵ for one channel of MetalMapper data, using the ISO polarizability estimate in figure 16. While this approach dictates dipolar noise, the actual realization of noise that produced the estimated (unsmoothed) polarizabilities is not necessarily dipolar - there may be another noise model that can produce the same result. Nonetheless, we can use the statistics of the predicted dipolar noise to understand the character of the noise that produced the observed polarizabilities. Repeating the polarizability fitting process for all ISO items at Pole Mountain, we obtain the noise distribution for a single channel in figure 18. The resulting distribution can be decribed as *mostly* normal: larger amplitude outliers produce heavier tails than would be obtained with a purely Gaussian noise distribution. This type of heavy-tailed noise often motivates the use of robust norms that downweight the influence of outliers on model estimates. We investigated robust fitting in an earlier project and showed that these methods are especially beneficial when processing dynamic data collected with monostatic EM sensors (Beran et al., 2011). In contrast, the large number of (nearly) redundant measurements acquired with multistatic sensors make these data inherently robust to outliers. Robust processing therefore offers negligible improvements in parameter estimation for these platforms.

Symmetric, heavy-tailed noise is often represented as a mixture of two Gaussian components with equal means (Marrona et al., 2006). The lower weight, larger variance component describes the outlying noise. In figure 18 we find that a single Gaussian does not adequately describe the noise, and so we use two techniques to fit univariate, Gaussian mixture models

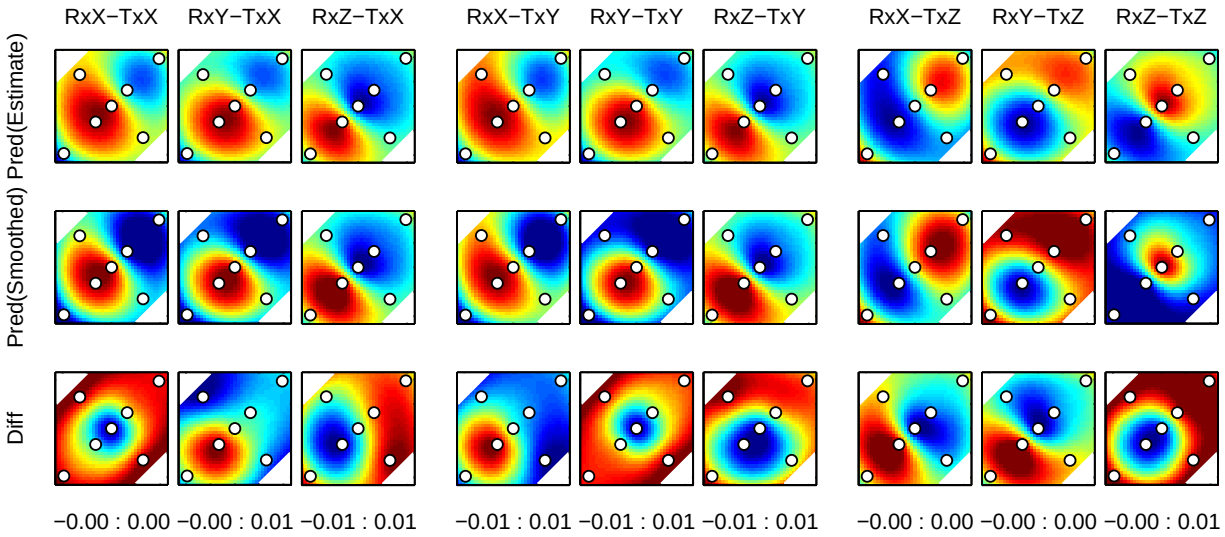


FIGURE 17. MetalMapper data at channel 42 (7.9 ms) predicted using unsmoothed (top row) and smoothed (middle row) polarizability estimates from figure 16. The difference between the predicted data sets (bottom row) is itself a dipolar anomaly and is an estimate of the noise realization on the data. Markers indicate receiver locations.

to the empirical distribution of the noise. Expectation-maximization (E-M) is a popular, iterative algorithm for maximum likelihood estimation of mixture model parameters (Hastie et al., 2001). We have also developed an alternative estimation method that minimizes the

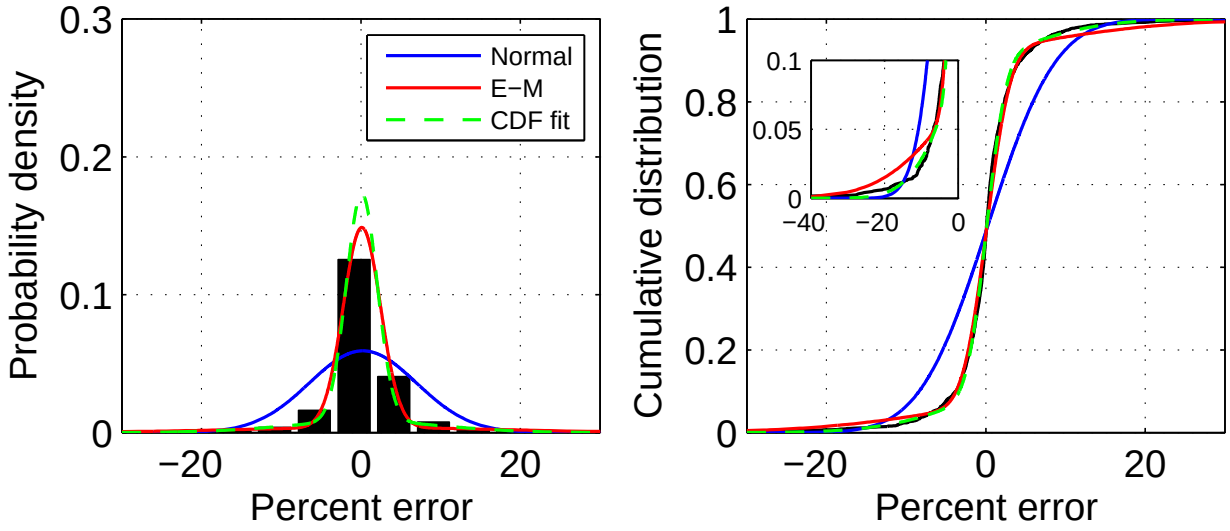


FIGURE 18. Fits to noise distribution derived from ISO target polarizabilities, for channel 42 of MetalMapper data at Pole Mountain. Left: Empirical density (histogram) and predicted densities using a single normal distribution (denoted Normal) and a mixture of normal distributions. The mixture models are estimated using expectation-maximization (E-M) and a novel method that directly fits the CDF. Right: empirical (black line) and predicted cumulative densities. Inset shows fits in the shoulder of the distribution.

least squares difference between empirical (observed) and predicted cumulative distribution functions (CDFs). By specifying particular points on the CDF that we want to fit, we place more emphasis on the "shoulders" of the distribution and obtain better agreement between observed and predicted cumulative distributions. Fitting the CDF directly is motivated by the fact that the Kolmogorov-Smirnov (K-S) hypothesis test of two non-parametric distributions uses the maximum difference between CDFs as its test statistic (Press et al., 1992). By this criterion, our method produces a higher likelihood estimate of mixture parameters than E-M. The CDF fitting method produces a good fit across all time channels, indicating that a two-component mixture is an accurate model for the noise.

Figure 19 shows the standard deviations and weights of the normal mixture components across all MetalMapper channels at Pole Mountain. The lowest noise variances are at intermediate channels. At late times the observed decays hit the background noise floor, so that the noise standard deviation is a larger proportion of the true datum.

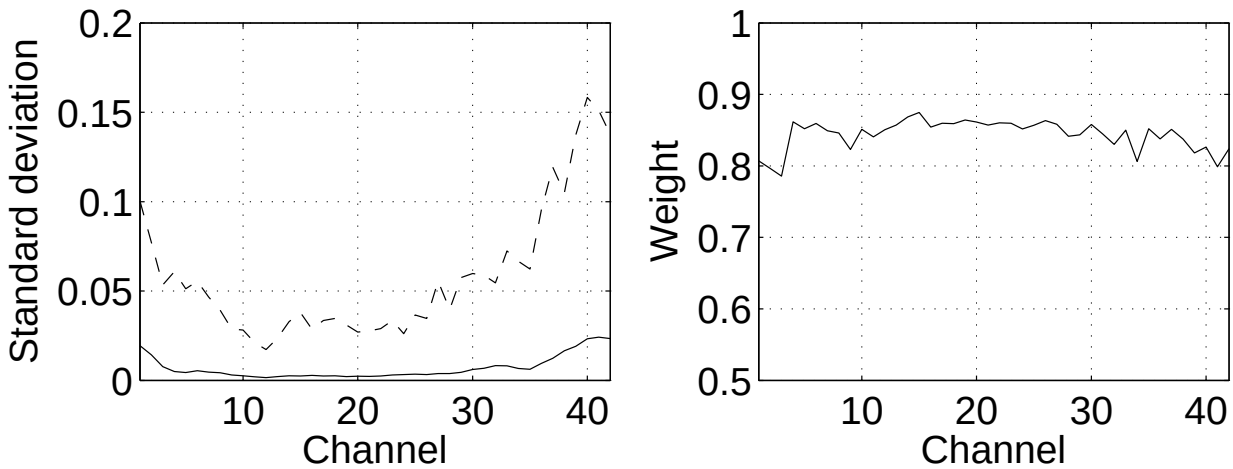


FIGURE 19. Estimated normal mixture model parameters at each MetalMapper channel. Left: Standard deviations of normal mixture components, expressed as a percentage of the observed datum, estimated via CDF fitting technique. Right: Weighting of first, smaller variance component in mixture model.

In this section we have investigated, in some detail, methods for estimating the distribution of additive noise that contaminates observed data. We conclude that a simple model of uncorrelated Gaussian noise is only roughly applicable to MetalMapper data. A mixture of correlated Gaussians is a much better approximation. With these lessons in mind, we now turn to a straightforward linear approach for modeling the effects of noise and predicting the distributions of estimated polarizabilities. This simple method makes no restrictive assumptions about the distribution of the noise and directly reproduces the variability of polarizabilities observed in field data.

4.1.2. *Predicting the distributions of polarizabilities.* Equation 23 relates the observed noise ϵ to the perturbation in the model $\Delta\mathbf{m}$ via the forward modeling operator $\mathbf{G}$

$$\epsilon = \mathbf{G}\Delta\mathbf{m}. \tag{23}$$

From this relation, we can express the covariance of the noise as

$$\begin{aligned} \mathrm{cov}(\epsilon) &= E(\epsilon\epsilon^T) \\ &= \mathbf{G}E((\Delta\mathbf{m})(\Delta\mathbf{m})^T)\mathbf{G}^T \\ &= \mathbf{G}\mathrm{cov}(\mathbf{m})\mathbf{G}^T. \end{aligned} \tag{24}$$

We can estimate the covariance of the polarizabilities, $\mathrm{cov}(\mathbf{m})$ directly from the set of estimated target polarizabilities at a site, as displayed, for example, in figure 12. Note that for each target in this set, there is a separate $\mathbf{G}$ corresponding to the estimated orientation and location of that target.

We now assume that there is some equivalent dipole source at the median location $\mathbf{r}_{eq}$ of all targets in our observed target set. For example, at Pole Mountain and Camp Beale the median location of ISOs is directly below the center of the array, at approximately 20 cm depth. The forward operator $\mathbf{G}_{eq}$ at this location can then be used to compute the effective noise covariance using equation 24.

$\mathbf{G}_{eq}$ also depends on the orientation of the dipole source. As in a GSV analysis, we can restrict ourselves to targets in a particular orientation. Alternatively, we can specify a probability distribution of target orientations and marginalize over azimuth and dip to determine the expected covariance of the estimated polarizabilities. Appendix B develops expressions for the covariance of the estimated polarizabilities under a uniform distribution of target dip and azimuth. For simplicity, in this section we consider only vertically-oriented targets.

Figure 20 compares the effective noise covariances for Pole Mountain and Camp Beale computed using this approach. Consistent with our characterizations of the noise in the previous section, there is significant off-diagonal structure in figure 20(a), indicating the presence of correlated noise on the data. The estimated standard deviations in figure 20(b) show that a single standard deviation at each time channel, as assumed in the previous section, is not an accurate characterization of the noise. Instead, we find that are three standard deviations required at each channel: one corresponding to maximum coupling between transmitter and receiver (e.g. x transmitter and x receiver component), and two components for perpendicular transmitter/receiver combinations.

Inverting at a fixed target location is a linear transformation of the data, and so with the noise covariance from equation 24, we can directly compute the covariance of the polarizabilities at the location $\mathbf{r}_j$ with corresponding forward operator $\mathbf{G}_j$ as

$$\mathrm{cov}(\mathbf{m})_j = \mathbf{G}_\mathbf{j}^\dagger \mathrm{cov}(\epsilon) \mathbf{G}_\mathbf{j}^{\dagger^T}, \tag{25}$$

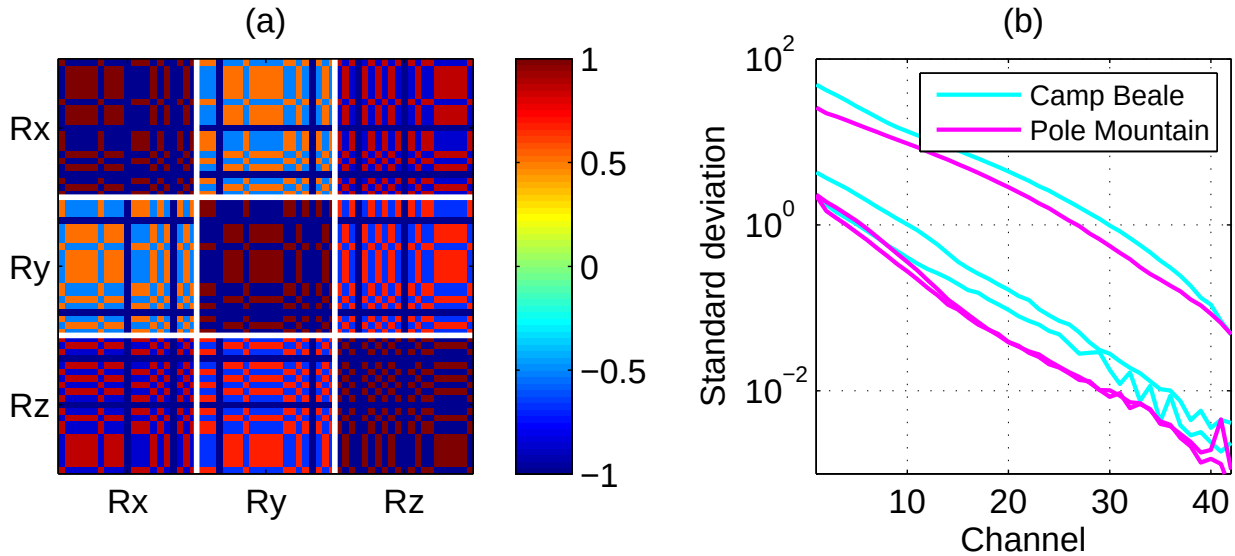


FIGURE 20. Structure of the noise covariance matrices at Camp Beale and Pole Mountain. (a) Correlation matrix structure. Each block in the correlation matrix is comprised of the 21 measurements (7 receivers $\times$ 3 transmitters) made with a given receiver component (b) Estimated standard deviation of the noise for center MetalMapper receiver (all components) at Camp Beale and Pole Mountain.

with $\mathbf{G}_j^\dagger$ denoting the pseudo-inverse. For $\mathbf{r}_j = \mathbf{r}_{eq}$, we find $\text{cov}(\mathbf{m})_j = \text{cov}(\mathbf{m})_{eq}$. This means that a target at our equivalent source location will produce a distribution of polarizabilities that is exactly that which we recovered from the field data.

Now as we vary the target location $\mathbf{r}_j$, there is a commensurate change in our model uncertainty, as shown in figure 21. This analysis is still restricted to vertically oriented targets. The change in model uncertainty reflects the relative change in curvature of the misfit function as the target location varies. Figure 22 shows the dependence of the polarizability uncertainty on the horizontal location of the target. The uncertainty is relatively constant within the sensor footprint.

At each possible location for a target, we now have the machinery to compute the uncertainty in the polarizabilities that is expected given a specified noise covariance. In practice, we have a sample of targets S at different locations and wish to compute the overall covariance of this sample. It is straightforward to show that the covariance of the sample can be estimated as the mean of the covariances of all targets in the sample

$$\text{cov}(\mathbf{m})_S = \frac{1}{M}\sum_j^M \text{cov}(\mathbf{m})_j. \tag{26}$$

More generally, for a specified distribution of target locations $p(\mathbf{r})$ the model covariance is computed via the expectation

$$\text{cov}(\mathbf{m}) = \int_{\mathbf{r}} \text{cov}(\mathbf{m})(\mathbf{r})p(\mathbf{r})d\mathbf{r}. \tag{27}$$

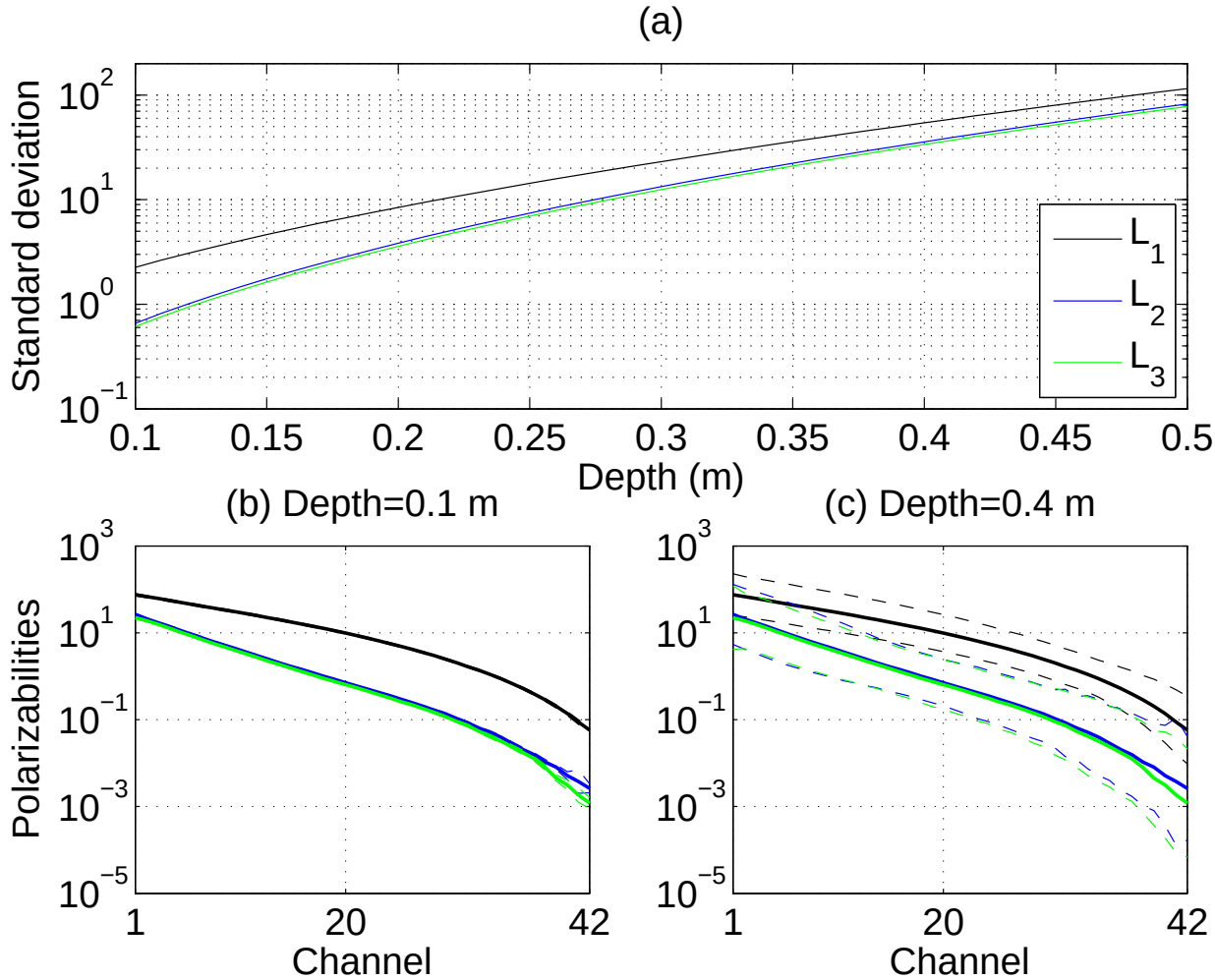


FIGURE 21. Effect of target depth on polarizability uncertainty. (a) Predicted standard deviation of polarizabilities as a function of target depth. Predictions are for channel one MetalMapper data at Pole Mountain. (b) ISO reference polarizability (solid lines) and predicted 95% confidence interval (dashed lines) for ISOs at 0.1 m depth. (c) As in (b), at 0.4 m depth.

For specified horizontal and vertical spatial distributions of targets, we approximate the above integral at a discrete set of points, with each point weighted by

$$P(\mathbf{r}_j) = \frac{p(\mathbf{r}_j,)}{\sum_{k=1}^{M} p(\mathbf{r}_k)}. \tag{28}$$

Figure 23 shows the effect of assuming normal or uniform horizontal distributions of ISO targets on the predicted polarizability distributions. The uniform horizontal distribution in (a) might represent a scenario where the operator of the sensor platform does a relatively poor job of positioning the sensor on top of a target. If we replace this operator with a more experienced/careful driver in (b), there is a moderate expected reduction in polarizability variance. In cases (a) and (b), we have assumed a uniform vertical distribution of ISOs down to a maximum depth of 40 cm[1]. Consequently, the overall, averaged variance

[1]This maximum depth may correspond to the clearance depth specified by a regulator.

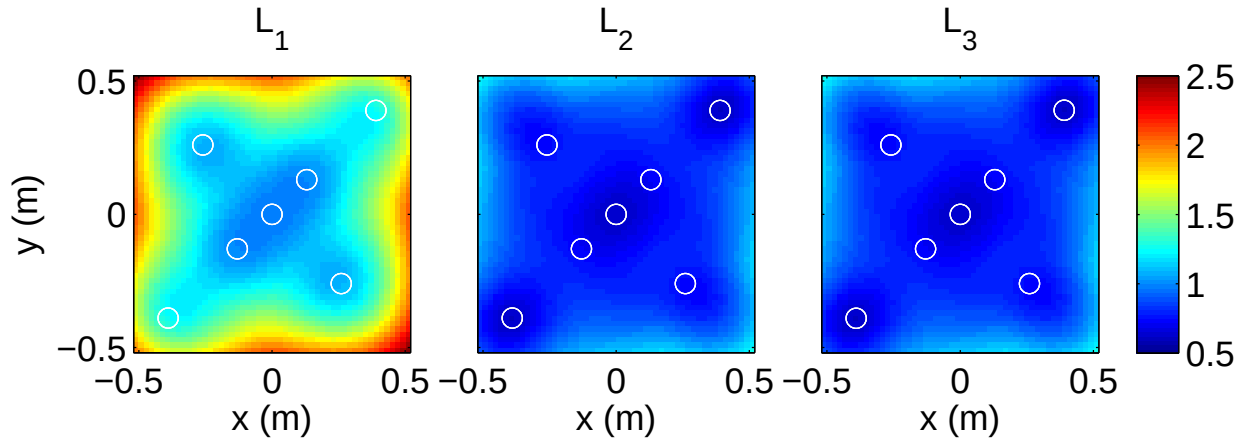


FIGURE 22. Dependence of polarizability standard deviations on horizontal target location, Gridded images show $\log_{10}$-transformed standard deviations. Markers indicate MetalMapper receiver locations. Calculations are for a target at 20 cm depth, using data covariance derived from Pole Mountain.

from equation 27 is dominated by targets at depth and the benefits of improved horizontal positioning are attenuated. In (c) and (d) we now reduce the maximum clearance depth for ISOs to 20 cm. This has a strong effect on the polarizability variance for both drivers, with a dramatic reduction in variance for our experienced driver in (d). We conclude from these simple examples that the important controlling variable in polarizability variance with a cued sensor is target depth. Good horizontal positioning is important, but may only provide modest improvements in parameter estimation when we are tasked with classifying targets near their maximum detection depth.

Thus far we have focused exclusively on MetalMapper data, but the technique is readily extended to any cued sensor, with the functionality to compute $\mathbf{G}$ implemented for any data type we already invert (TEMTADS2x2, MPV, etc.). Figure 24 compares recovered ISO polarizability distributions for the BUDhh and TEMTADS2x2 sensors deployed at Camp Beale. The BUDhh measures data over a very short time window, with the result that there is very little random jitter in the recovered polarizabilities, relative to the late time polarizabilities measured with the TEMTADS2x2. Similar to its larger 5x5 antecedent, the noise on the 2x2 array could be substantially reduced by simply averaging the polarizabilities in windows of e.g. 5 channels. Evident for both sensors are correlated shifts of the polarizabilities, in this case likely due to a strong background response at the site.

We use these polarizability distributions to generate noise covariances for each sensor, and then to predict the distributions of polarizabilities that would arise for each sensor for ISOs distributed uniformly within each sensor's footprint. We again consider how different target distributions with depth affect polarizability variance. In the first scenario (middle row of figure 24), we represent the distribution of ISOs at Camp Beale as a normal distribution with mean depth 15 cm and a standard deviation of 5 cm. This roughly corresponds to the actual depth distribution of ISOs at Beale. Unsurprisingly, this model closely reproduces

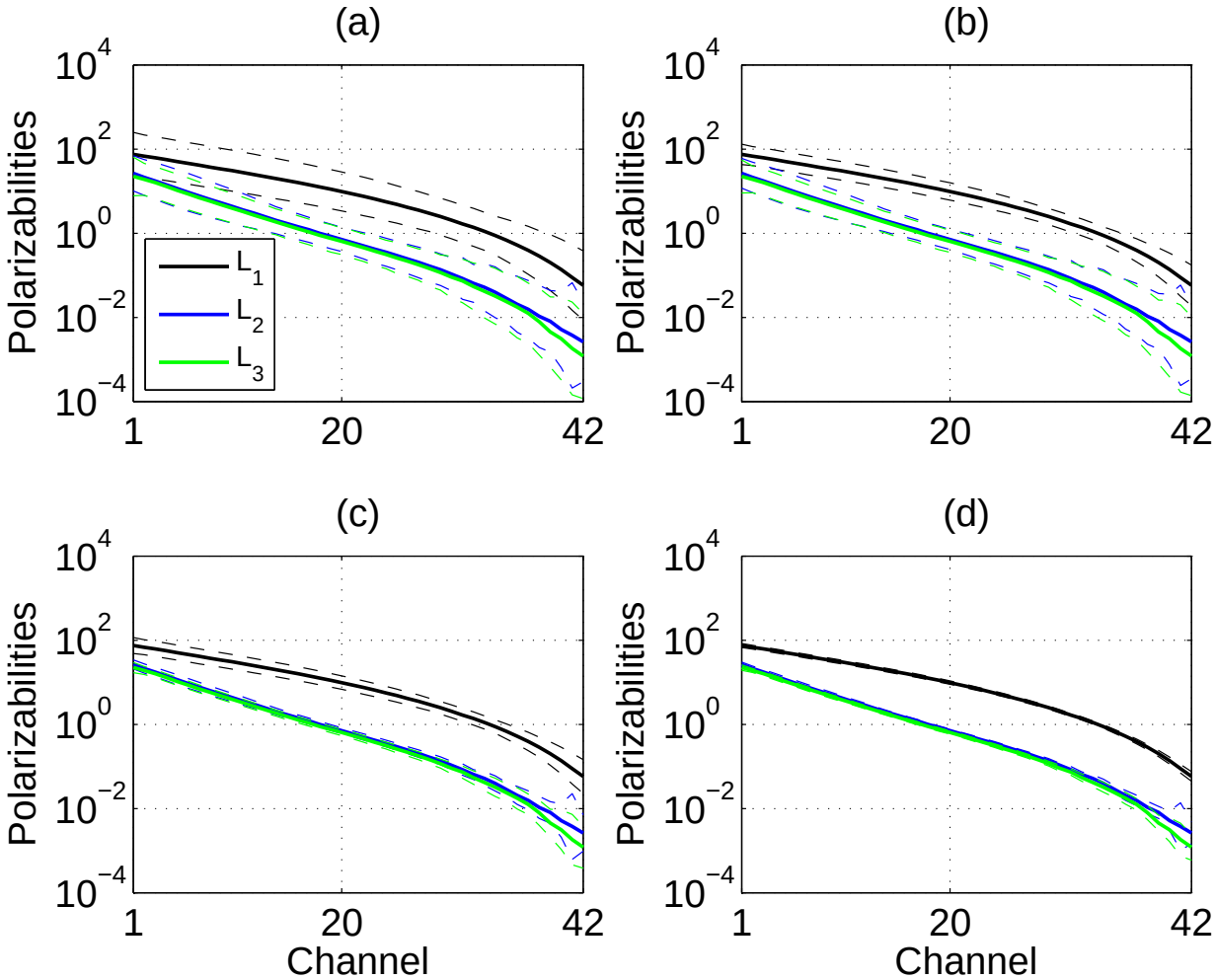


FIGURE 23. Predicted ISO polarizabilities for four spatial distribution scenarios with Pole Mountain noise model. (a) Uniform horizontal distribution across the sensor footprint, uniform vertical distribution between 0.05 and 0.4 m. (b) Normal distribution truncated on the horizontal interval $x, y \in [-0.3\ 0.3]$ m, with standard deviation of 0.1 m in both x and y. Uniform vertical distribution between 0.05 and 0.4 m depth. (c) As in (a), but with maximum vertical depth constrained to 0.2 m. (d) As in (b), but with maximum vertical depth constrained to 0.2 m.

the observed polarizability distributions for both sensors. Now if we make the parameter estimation problem a little bit harder by distributing targets uniformly with depth down to 30 cm (bottom row of figure 24), the BUDhh shows a significant increase in polarizability variance, relative to the 2x2. This is likely because the BUDhh measures fewer data at each sounding (30 data at each channel) than the 2x2 (48 data at each channel). This produces a larger rate of increase in the polarizability variance with depth for the BUDhh than the TEMTADS2x2.

In the preceding discussion we have used a straightforward linear analysis to develop methods for predicting the distribution of estimated polarizabilities given site specific noise. Beyond its simplicity, the approach is appealing in that it combines the problem physics -

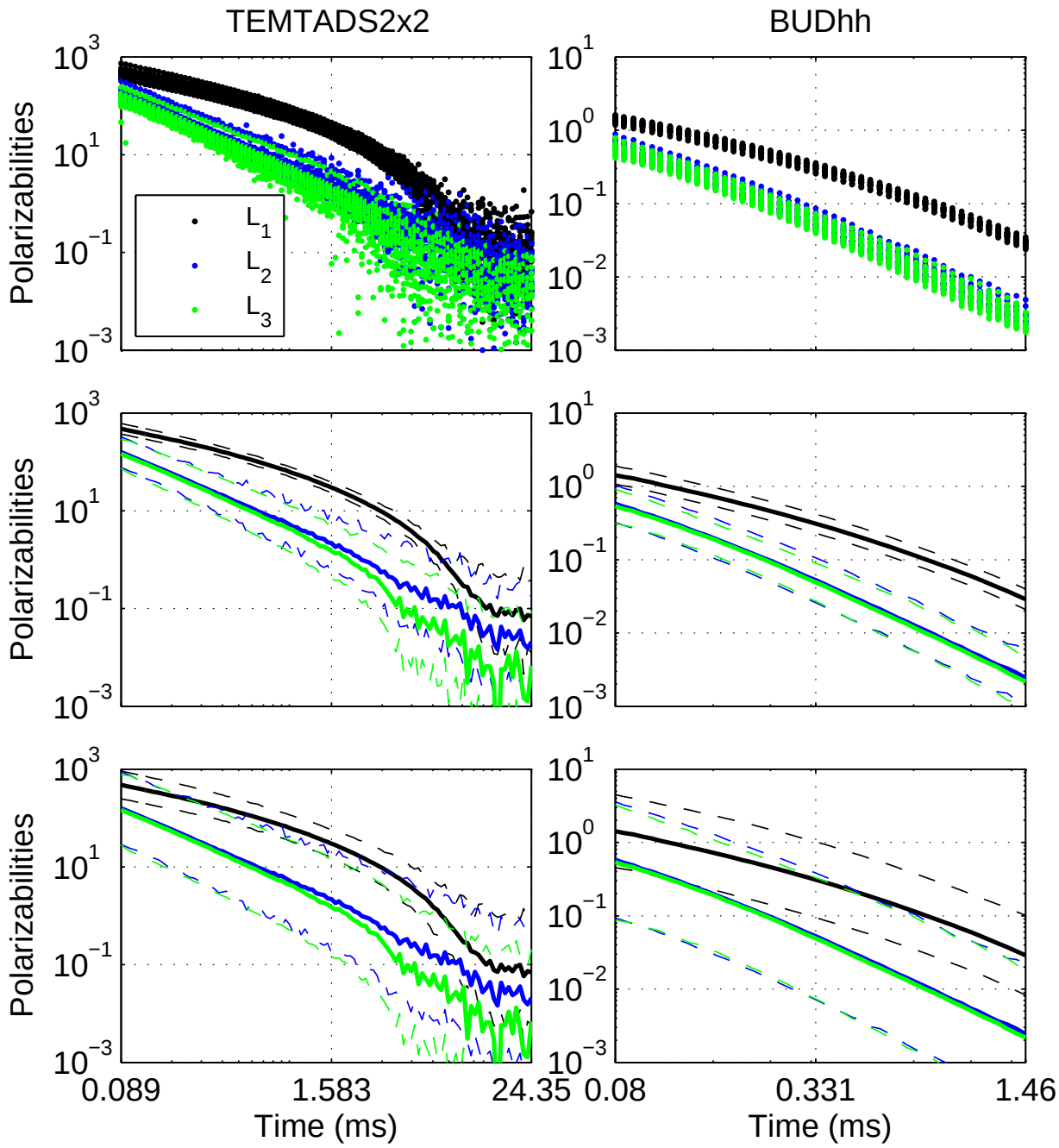


FIGURE 24. ISO polarizabilities for TEMTADS2x2 and BUDhh handheld sensors at Camp Beale. Top row: actual ISO polarizabilities for each sensor at Camp Beale. Middle row: Predicted ISO polarizability distributions for each sensor with a uniform horizontal distribution of targets within the sensor foorprint. Targets are normally distributed in depth, with a mean depth of 15 cm, and 5 cm standard deviation. Bottom row: As in middle row, but with targets uniformly distributed with depth down to 30 cm.

encapsulated in the forward modeling matrix $\mathbf{G}$ - with a realistic noise model described by a dense noise covariance.

In a linear uncertainty analysis we typically assume perfect recovery of target location. There is, however, some accounting for imperfect location recovery, and other nonlinear effects, implicit in our estimation of the noise covariance. This is because we use a set of actual, estimated polarizabilities for a target class to back out the noise covariance. Correlated shifts of these recovered polarizabilities resulting from errors in location estimation are therefore propagated through the analysis. We can further quantify the magnitude of uncertainty in position estimation at each site via a linearized apparaisal. This is identical to the linear uncertainty analysis, except the forward matrix $\mathbf{G}$ in equation 25 is replaced by the sensitivity (Jacobian) matrix $\mathbf{J}$ of first derivatives with respect to target position. Figure 25 shows this analysis for the MetalMapper at Camp Beale and Pole Mountain. At each target location we compute the standard deviation of our location estimate as

$$\sigma_r = \sqrt{trace(\mathrm{cov}(r_i, r_j))} \tag{29}$$

with $\mathrm{cov}(r_i, r_j)$ denoting the covariance between i^{th} and j^{th} elements of the position vector $\mathbf{r}$. In figure 25(a) and (c) we show the uncertainty in target position at 30 cm depth for a single object scenario. Within the sensor footprint the median standard deviation in target location is approximately 30 cm and 15 cm at Camp Beale and Pole Mountain, respectively. In the linearized analysis we can also examine the effect of multi-object scenarios on location uncertainty. In (b) and (d) we introduce a second target near the edge of the array and compute the standard deviation in position as a function of the location of the first target. If the sensor is properly centered over one of the two targets, then the uncertainty is essentially unchanged. It is only when the two targets are both near an edge of the array that we expect a large uncertainty in recovered target location. Improved infield QC procedures should prevent this scenario, so that we can reasonably expect a relatively small error in recovered target location with cued sensors.

Thus far, our analysis is tied to specific sites for which we have a representative sample of ISO items to characterize model variability. Ultimately our goal is to predict classification performance for an arbitrary site characterized by particular conditions (topography, target density, etc.) If, for the moment, we regard Pole Mountain and Camp Beale as end members in the spectrum of classification difficulty, then we might envision a system where a user might select a reference site most similar to their problem, and then use the approach developed here to predict polarizabilities and classification performance. As more data sets with seeded ISOs become available, we will obtain a more diverse set of site conditions, with some even more challenging that Camp Beale (e.g. the planned demonstration in Waikoloa, HI). In ongoing work we will work to develop metrics (target density, magnitude of soil response, sensor positioning, etc.) necessary to characterize each site. We will also explore ways to average existing noise models (i.e. covariances) to make predictions for novel site conditions.

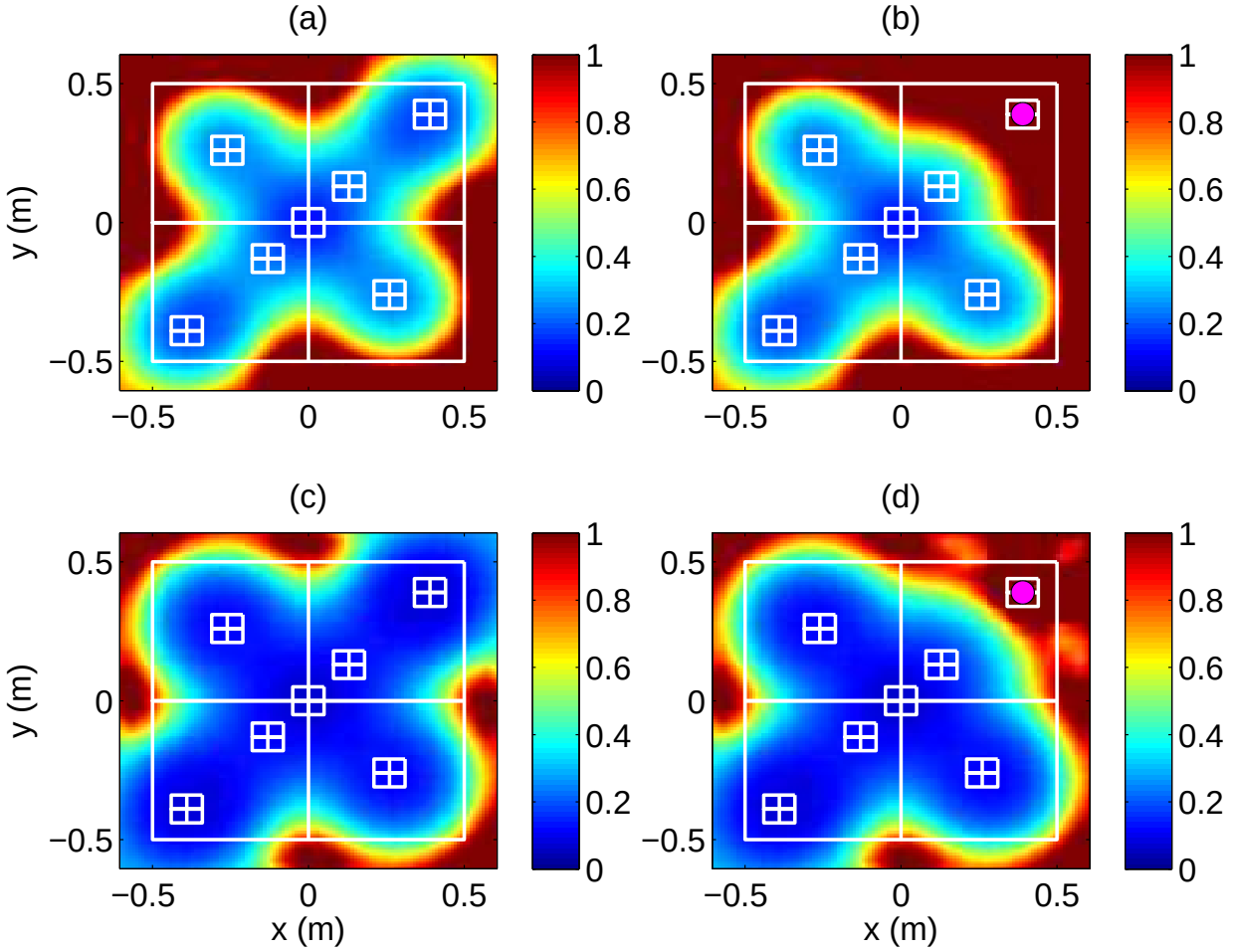


FIGURE 25. Standard deviation of positional error as a function of target position. (a),(c) Single-object scenario at Camp Beale and Pole Mountain, respectively. (b),(d) Two-object scenario at Camp Beale and Pole Mountain, respectively. Marker indicates fixed location of second target in top right corner of the array. Gridded image then shows the uncertainty in the location estimate for the first target.

4.1.3. *Predicting the ROC.* The next step in predicting classification performance is to use the distributions of polarizabilities to compute the distribution of the decision statistic for each target class. At this point we make the assumption that estimated polarizabilities are lognormally distributed, with variances predicted using the above analysis. This ensures that the polarizability distribution is non-negative. The log-transformed polarizabilities are then, conveniently, normally-distributed, allowing us to analytically compute the distribution of the decision statistic. For classification with multistatic sensor data, the decision statistic ϕ is typically a misfit of log-transformed estimated polarizabilities L^{est} with respect to some reference, or library, polarizability L^{ref}

$$\begin{aligned} \phi &= \sum_{i=1}^{N} (\log(L_i^{est}) - \log(L_i^{ref}))^2 \\ &= \sum_{i=1}^{N} X_i^2 \end{aligned} \tag{30}$$

with $X_i = \log(L_i^{est}) - \log(L_i^{ref})$. The summation extends over some subset of the polarizabilities. For high SNR targets we often use all three principal polarizabilities, whereas for noisier cases we might only use the primary polarizability to compute the decision statistic and rank targets. Other similarity measures may also be used as a decision statistic for ranking targets. For example, we often use a heuristic function of the form

$$\phi = \sum_{i=1}^{N} \left(\frac{(L_i^{est})^\gamma - (L_i^{ref})^\gamma}{1/N \sum_{j=1}^{N} (L_j^{ref})^\gamma} \right)^2 \tag{31}$$

with $\gamma \approx 0.1$. The following analysis can readily be applied to nonlinear functions of this form; we can use a first order approximation to compute the variance of the polarizabilities under arbitrary nonlinear transformations.

Continuing with the decision statistic in equation 30, we require the expected value of the log transformed polarizabilities in a class to be the logarithm of the true, reference polarizabilities of that class

$$E(\log(L_i^{est})) = \log(L_i^{ref}). \tag{32}$$

If the lognormally-distributed polarizability has predicted variance $\mathrm{var}(L)$, then the variance of the corresponding log-transformed variable is given by

$$\mathrm{var}(\log(L)) = \log \left(\frac{1 + \sqrt{1 + 4\mathrm{var}(L) \exp(-2E(\log(L)))}}{2} \right) \tag{33}$$

The mean and variance of the decision statistic are

$$
\begin{aligned}
E(\phi) &= \sum_{i=1}^{N} E(X_i^2) \\
\mathrm{var}(\phi) &= \sum_{i=1}^{N} \mathrm{var}(X_i^2).
\end{aligned}
\tag{34}
$$

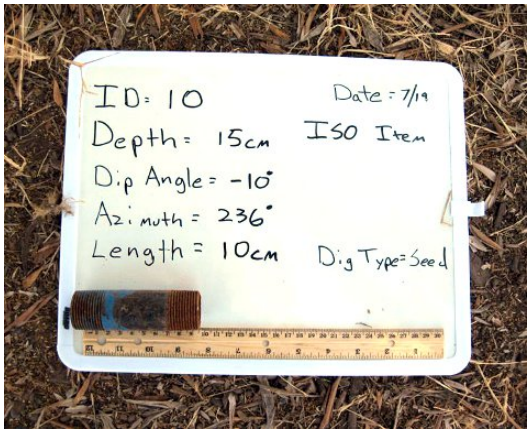
For the individual X_i, the required moments are

$$
\begin{aligned}
E(X_i^2) &= \mathrm{var}(X_i) + E(X_i)^2 \\
\mathrm{var}(X_i^2) &= E(X_i)^4 + 6E(X_i)^2\mathrm{var}(X_i) + 3\mathrm{var}(X_i)^2 - E(X_i^2)^2
\end{aligned}
\tag{35}
$$

The above expressions can be used to compute the moments of the polarizability misfit for a specified target class with respect to an arbitrary reference polarizability. The sum of N standardized normal random variables is distributed as a χ^2 random variable with N degrees of freedom. In this case, the X_i are not standardized (i.e. normalized by their standard deviation). We therefore use a variant of the central limit theorem: the summation of independent, but not identically-distributed, variables in equation 30 will tend to a normal distribution (Billingsley, 1995) with mean and variance given by equation 34.

As an example, we consider discrimination between ISO targets and a clutter item. Ground truth photos are shown in figure 26. We have selected a non-TOI item from Camp Beale that has a very similar primary polarizability to an ISO. In figure 27 we simulate the distributions of polarizabilities that would be obtained for these items assuming the Pole Mountain noise covariance, and a uniform spatial distribution for both targets to a maximum depth of 40 cm. We first compute the decision statistic using only the primary polarizability (top row of figure 27). This produces some overlap in the distributions of ϕ, such that the resulting ROC requires us to dig approximately 5% of clutter in order to ensure that 99% of all TOI are found. If we instead use all polarizabilities for classification (bottom row of figure 27), then the ROC indicates perfect classification. This is because there is a significant difference between transverse polarizabilities for these targets at the specified noise level.

Figure 28 considers the same classification problem as in figure 27, but uses the noise covariance derived from Camp Beale. The increased noise at this site degrades classification performance using the primary polarizability, but perfect classification is still expected when all three polarizabilities are used to rank targets.



(a) ISO

(b) Clutter, Camp Beale target 2530

FIGURE 26. Targets used for ROC prediction.

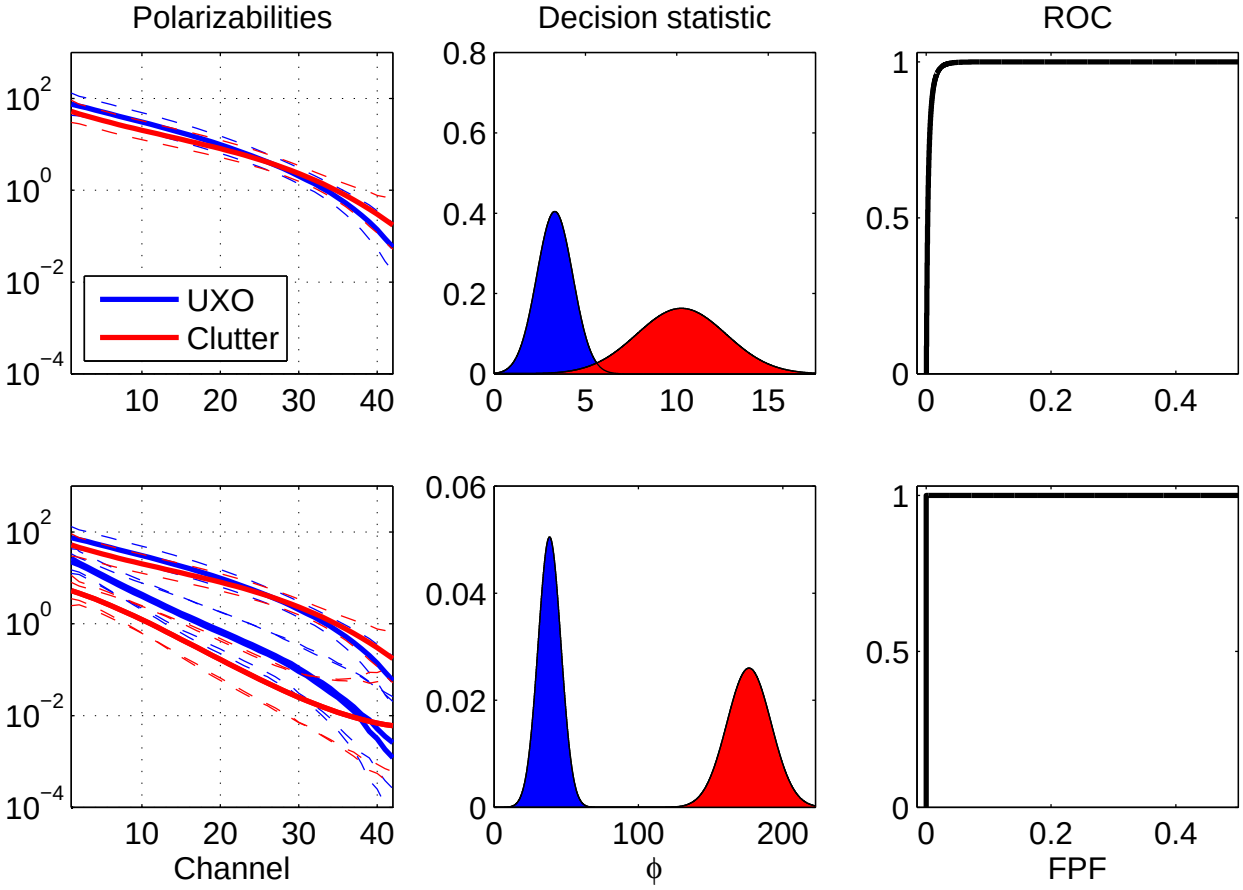


FIGURE 27. ROC prediction for the MetalMapper, Pole Mountain noise. In this scenario we discriminate between ISOs and a clutter item with a similar primary polarizability. Top row: Classification using primary polarizabilities only. Bottom row: Classification using all polarizabilities.

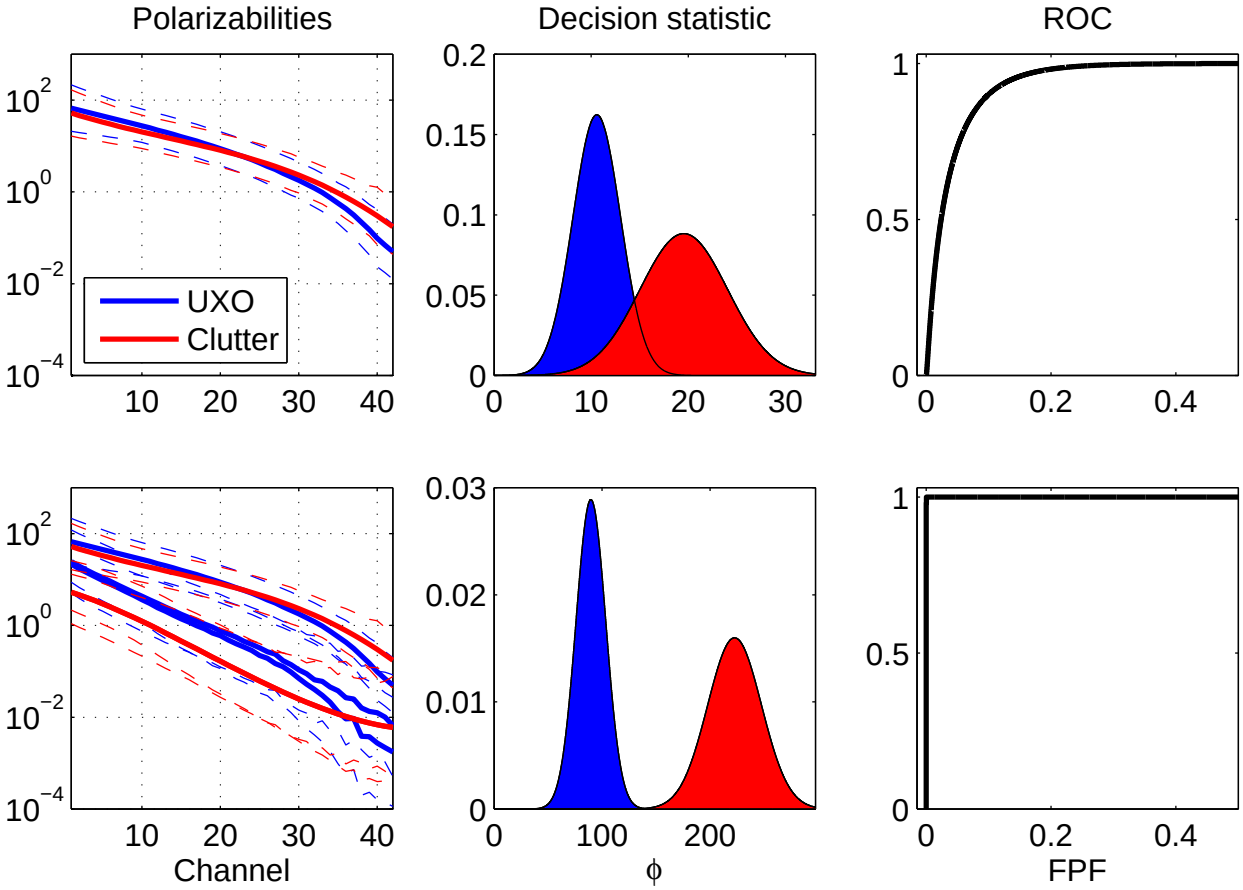


FIGURE 28. ROC prediction for the MetalMapper, Camp Beale noise. Plots are the same as in figure 27, only the noise covariance is changed.

Thus far, we have restricted our performance prediction analysis to targets in a single orientation, with the results in this section generated for horizontally-oriented targets. Figure 29 compares ROC prediction for the same TOI and non-TOI items in horizontal and vertical orientations, as well as for a uniform angular distribution of target orientations. Expressions for this latter case are developed in appendix B. Horizontal targets are the most

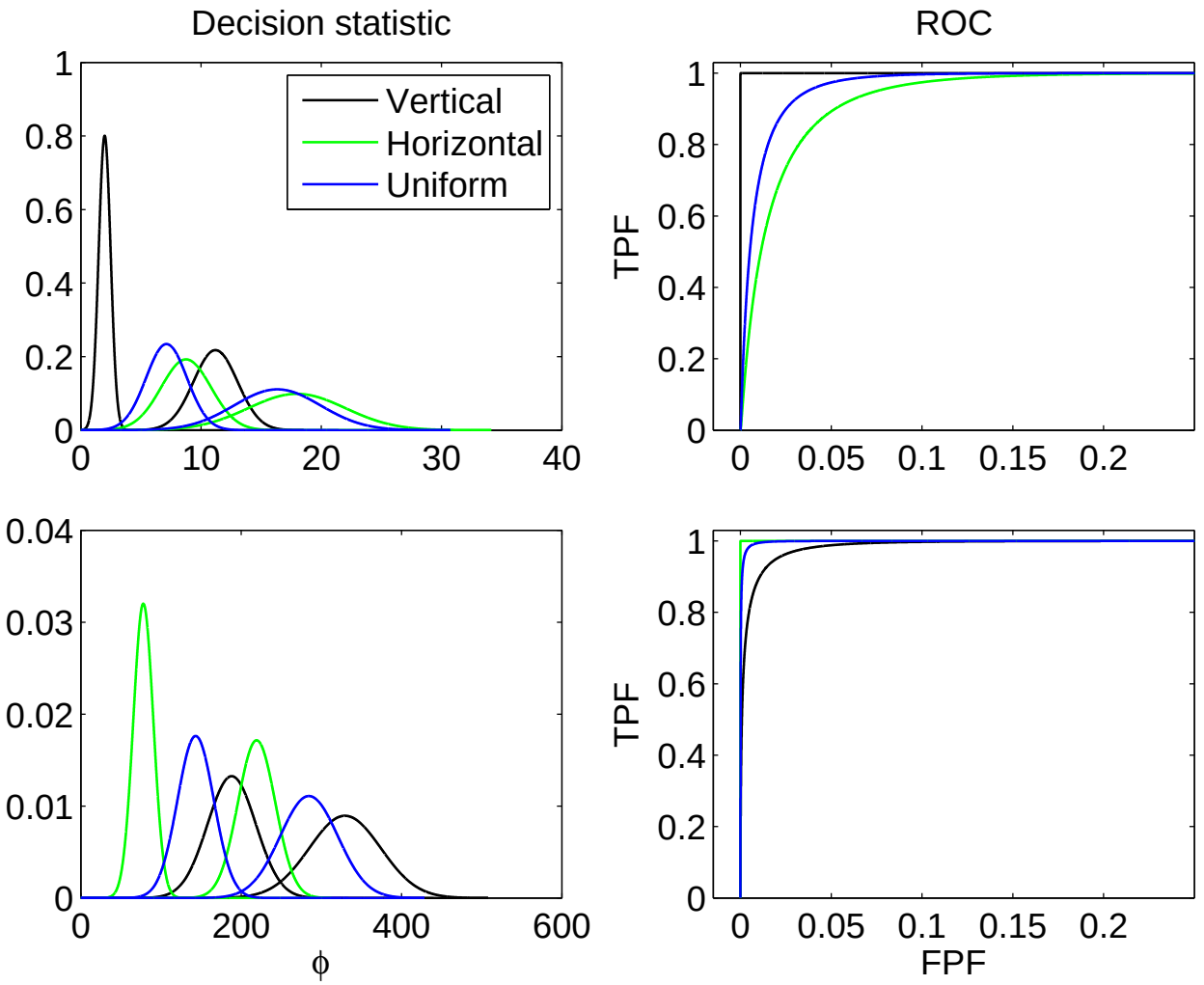


FIGURE 29. Effect of target orientation on decision statistic and ROC. Top row: decision statistic and ROC for classification with primary polarizabilities. Bottom row: decision statistic and ROC for classification with all polarizabilities.

difficult scenario for discrimination using the primary polarizabilities. In this orientation, the primary polarizability is less coupled to excitation by the horizontal MetalMapper coil. This is alleviated somewhat by the (horizontal) fields transmitted by the vertical x and y coils, but since the center of these loops is vertically displaced from the receiver plane, data from the vertical transmitter coils is somewhat lower SNR than from the horizontal coil. This results in more variability in the estimated primary polarizability for horizontally-oriented targets than for vertically-oriented targets. We verify this effect with numerical simulations in figure 30. We conclude that horizontal targets represent the worst case scenario for

both detection and classification (when we rely exclusively on primary polarizabilities for classification).

Unsurprisingly, a uniform angular distribution of targets gives a classification result that is intermediate to purely horizontal and vertical cases. The distributions of the decision statistic, using only the primary polarizability, are quite similar for horizontal and uniform cases. This is expected: a uniform distribution with respect to target dip is biased towards horizontal orientations[2].

Moving now to discrimination with all polarizabilities (bottom row of figure 29), the performance of vertical and horizontal cases is reversed. Secondary and tertiary polarizabilities are better constrained in horizontal orientations, and so inclusion of these features allows for perfect discrimination between the two targets considered in this example. Conversely, it is harder to excite the transverse response of a vertical target, and using transverse polarizabilities for classification of vertical targets therefore degrades performance. The uniform angular distribution is again intermediate to vertical and horizontal cases.

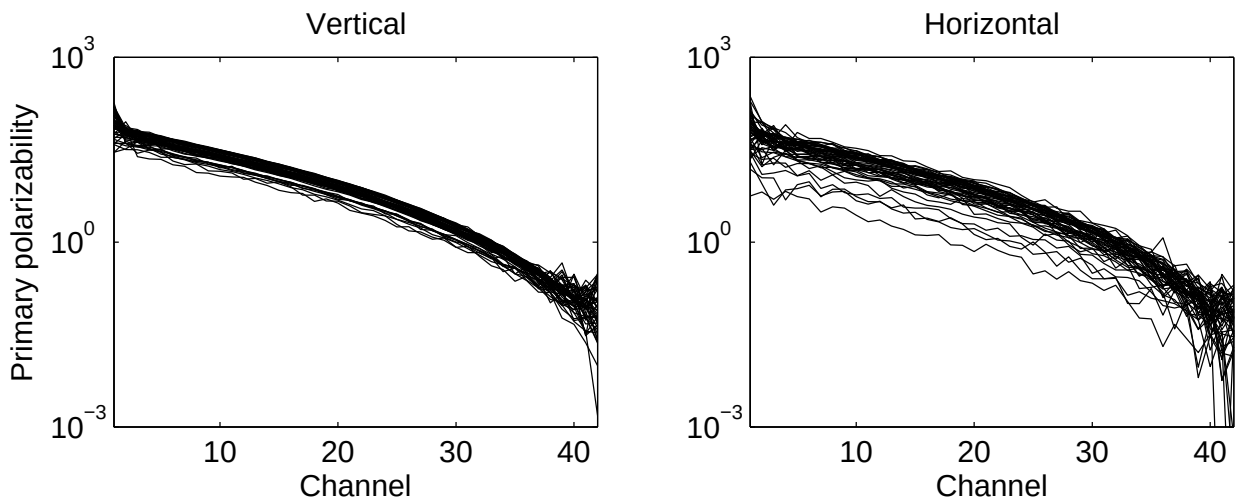


FIGURE 30. Numerical simulations of primary polarizability estimation for vertical and horizontal ISO targets. Targets are distributed on a uniform grid within the sensor footprint down to a maximum depth of 40 cm. Independent Gaussian noise is added at each channel using standard deviations estimated from Pole Mountain background measurements. The synthetic data for each location and noise realization are then inverted to produce the polarizability estimates.

In the preceding analyses, we have considered a simple classification problem: we have one set of TOI polarizabilities and one set of non-TOI polarizabilities. We then predict the ROC that would be generated for a given spatial and angular distribution of each target. By predicting classification performance for specific items, this analysis can give a site manager a concrete sense of the classification capabilities of a sensor under prescribed conditions.

[2]At a horizontal dip angle, an azimuthal angle g subtends a longer arc than at near vertical angles, so that a uniform distribution on the surface of a sphere produces more samples at the equator than at the pole.

A more general and realistic classification problem is to predict classification between a number of TOI classes and a prescribed set of non-TOI items. This is a straightforward generalization of the approach developed here: we compute the distribution of the decision statistic for each item, and then represent the distribution of the decision statistic within each class (TOI and non-TOI) as a mixture of the distributions of each item in that class. The component distributions in the sum can be weighted by probabilities reflecting the relative frequencies of each item. Alternatively, we can treat the non-TOI class as fundamentally uncertain and define an underlying non-TOI distribution. The predicted polarizability distribution arising for specified parameters (spatial and angular distributions and noise covariance) can then be convolved with the underlying distribution of non-TOI polarizabilities.

The choice of representative non-TOI items is an important consideration when predicting classification performance. We have maintained a database of targets from past ESTCP demonstrations, and we envision a decision support system with access to this database. Non-TOI could be selected based on visual similarity to TOI, or based on the similarity of polarizabilities to reference polarizabilities.

4.2. **Performance prediction for dynamic sensors.** In order for the performance prediction tools developed in the previous section to be useful for site managers and field geophysicists, a similar analysis must be available for dynamic sensors. We must be able to objectively show that the additional time and expense spent on collecting and processing cued interrogation data will provide tangible improvements in classification performance, and an overall reduction in cost. Our primary focus for this analysis is on EM-61 data, since this sensor is familiar to practicing geophysicists and provides a baseline performance metric. The approach can be readily generalized to other dynamic sensors, including arrays of EM-61s and newer dynamic platforms.

As with cued sensors, we can estimate a site-specific noise covariance using a linear transformation of the polarizabilities for ISO items. Figure 31 shows total polarizabilities for ISOs estimated from EM-61 cart data acquired at Camp Beale. These features exemplify a

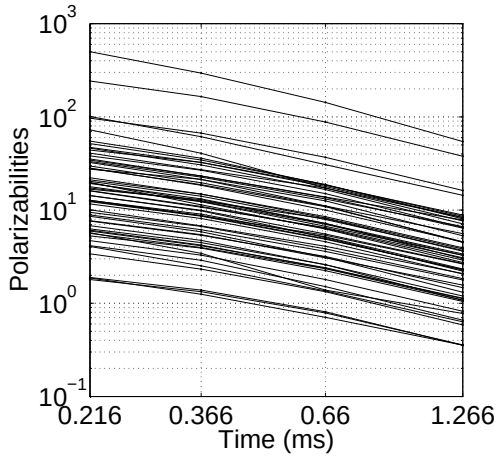


FIGURE 31. Total polarizabilities estimated for ISO targets at Camp Beale.

notorious difficulty with monostatic data processing: poor constraints on target depth produce a large spread in the amplitude of recovered polarizabilities. This is further illustrated in the size-decay feature space for the Camp Beale handheld area in figure 32. The poor constraint on estimated depth is manifested by the spread of the size parameter within each target class. The decay parameter, however, is relatively well constrained. In particular, small ISOs have a fairly tight distribution with respect to decay. This is also evident in figure 31: the primary polarizabilities have a reasonably consistent slope. We have found that the simplest and most reliable approach for classification with EM-61 data is a threshold on the decay parameter.

Using the approach developed in the previous section, we can predict the distribution of estimated polarizabilities for a given spatial distribution of targets. In this case we only specify the vertical distribution of targets, the horizontal distribution of targets will always be assumed uniform across the dynamic sensor's swath. However, a new complication arises when we consider dynamic data: the dimension of the forward matrix $\mathbf{G}$ depends upon the along-track and cross-track data density. This means that we cannot directly compute the

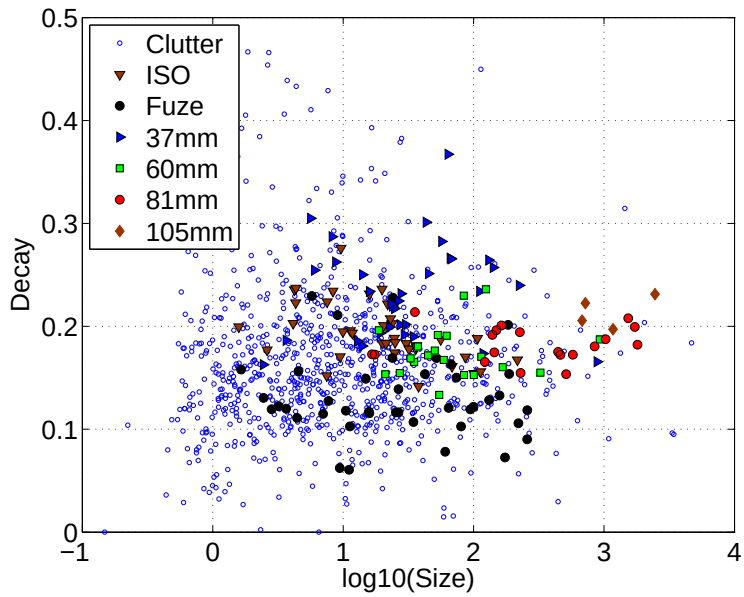


FIGURE 32. EM-61 size-decay feature space, Camp Beale handheld area.

effect of varying data density on polarizability variance, since the dimension of the estimated data covariance at a given density will not match the dimension of $\mathbf{G}$ for a different density.

To address this problem, we first associate the observed covariance of the polarizabilities $\text{cov}(\mathbf{m})$ within a target class with the forward modeling $\mathbf{G}_1$. Matrix $\mathbf{G}_1$ is evaluated at a median target depth, with a specified target orientation (horizontal or vertical), and a nominal along-track and cross-track data spacing. The model covariance at this location, without scaling by the data covariance, is

$$S_1 = (\mathbf{G}_1^T \mathbf{G}_1)^{-1}. \tag{36}$$

We denote the diagonal elements of $\mathbf{S}_1$ - the unscaled variances of the polarizabilities - as v_1. For $\mathbf{G}_2$ evaluated for a different data spacing (at the same target depth and orientation), we can similarly generate v_2. We then compute the ratio of standard deviations at the two data densities

$$\rho_i = \sqrt{\frac{v_{2i}}{v_{1i}}}. \tag{37}$$

Finally, we form the diagonal matrix

$$\mathbf{R} = diag(\rho). \tag{38}$$

and scale the model covariance of the target class by $\mathbf{R}$ to obtain the expected model covariance at the data density corresponding to $\mathbf{G}_2$

$$\text{cov}(\mathbf{m})_2 = \mathbf{R}\text{cov}(\mathbf{m})\mathbf{R}^T. \tag{39}$$

This transformation preserves the correlation structure of $\text{cov}(\mathbf{m})$ and scales the model covariance by the change in polarizability uncertainty dictated by the change in data density.

Figure 33 illustrates this effect for varying cross-track line spacing: as the line spacing is increased, the uncertainty in the total polarizability increases nonlinearly.

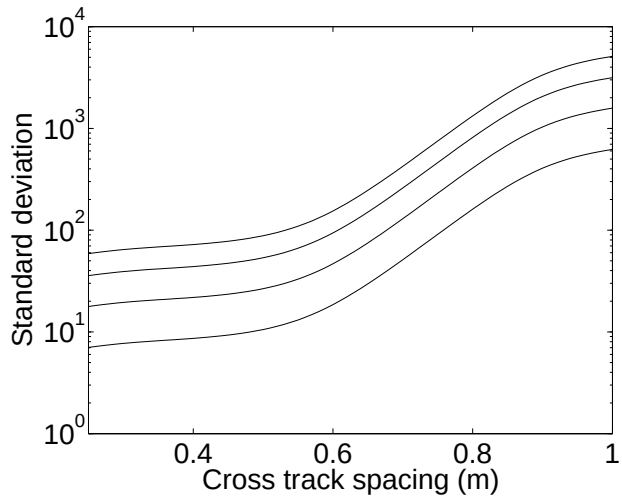


FIGURE 33. Dependence of EM61 total polarizability standard deviations at Camp Beale on cross track spacing. Standard deviations are displayed at all four EM61 channels, with the magnitude of the uncertainty decreasing with time.

We now turn to computation of the distribution of the decay parameter. The total polarizability at each time channel has variance

$$\mathrm{var}(L_{total}(t_k)) = \sum_{i,j} \mathrm{cov}(L_i(t_k), L_j(t_k)). \tag{40}$$

Treating the total polarizability as a lognormally distributed random variable, the decay parameter

$$\mathrm{decay}(t_k, t_j) = \frac{L_{total}(t_k)}{L_{total}(t_j)}, \tag{41}$$

is then the ratio of two lognormally distributed variables, which we now denote as X/Y. Taking the log of the ratio we have

$$Z = \log\left(\frac{X}{Y}\right) = \log(X) - \log(Y). \tag{42}$$

The log transformations of X and Y are normally distributed, so that Z is itself normally distributed with mean and variance

$$\begin{aligned} E(Z) &= E(\log(X)) - E(\log(Y)) \\ \mathrm{var}(Z) &= \mathrm{var}(\log(X)) + \mathrm{var}(\log(Y)) - 2\mathrm{cov}(\log(X), \log(Y)). \end{aligned} \tag{43}$$

The ratio $X/Y = \exp(Z)$ is then lognormally distributed with mean and variance

$$\begin{aligned} E(X/Y) &= \exp(E(Z) + \mathrm{var}(Z)) \\ \mathrm{var}(X/Y) &= (\exp(\mathrm{var}(Z)) - 1)\exp(2E(Z) + \mathrm{var}(Z)). \end{aligned} \tag{44}$$

The predicted distribution of the decay parameter depends on the variance of the total polarizabilities at the selected channels t_k and t_j. For EM61 data the decay parameter is usually computed at first ($t_j = 0.216$ ms) and fourth ($t_k = 1.225$ ms) time channels. We also find that the correlation between the total polarizabilities, quantified by the covariance term in equation 43, affects the decay distribution. Finally, the above transformations introduce a dependence on the size of the target, so that two items with the same expected decay parameter but different total polarizability amplitudes will produce different distributions of the decay parameter. These effects are illustrated in figure 34.

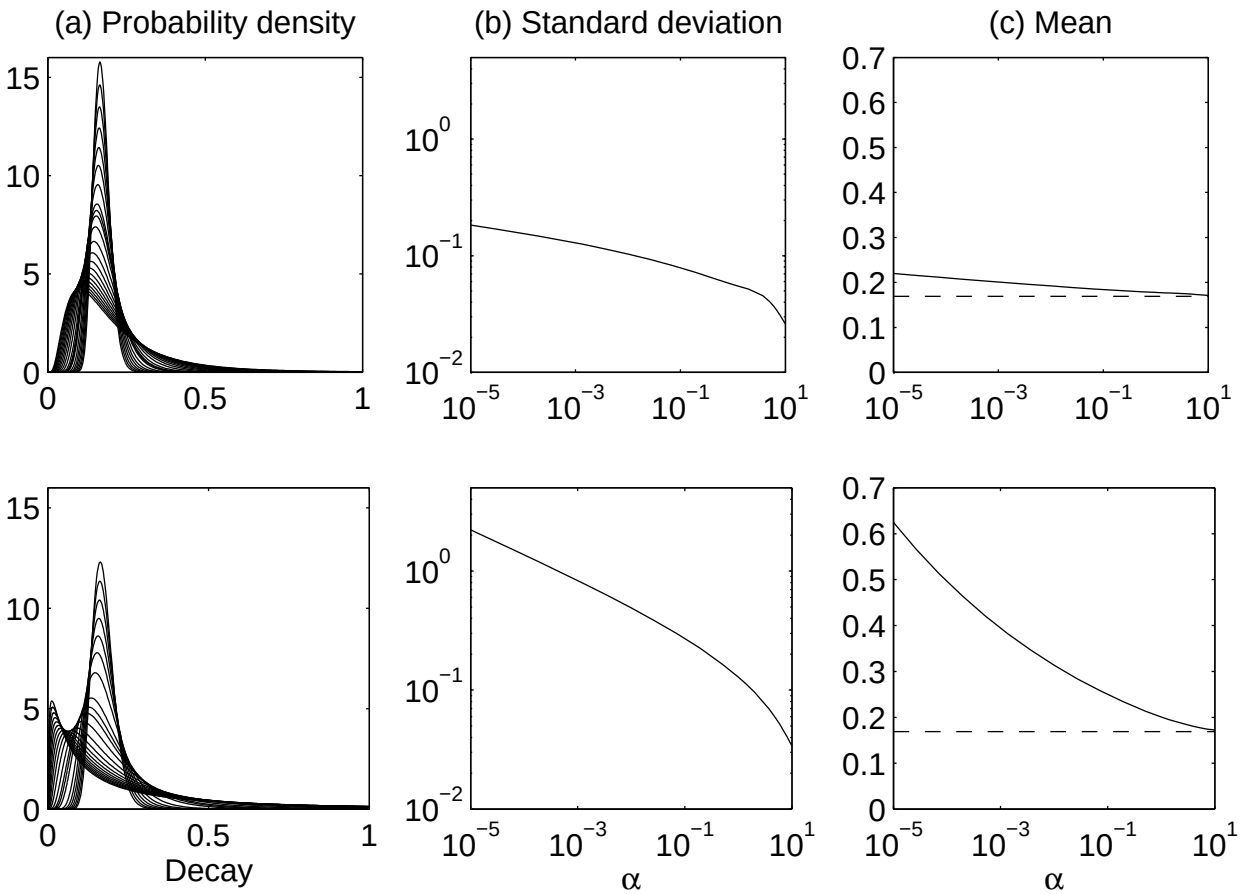


FIGURE 34. Dependence of lognormal decay distribution on target size and correlation coefficient. Top row and bottom rows are for correlations of $R = 0.99$ and $R = 0.95$, respectively. (a) Dependence of lognormal decay distribution on target size. As target size decreases, the mode of the decay parameter distribution shifts to smaller values. (b) Standard deviation of decay parameter as a function of size scaling parameter α. (b) Expected value of the decay parameter as a function of size scaling parameter α. Dashed line indicates the true value of the decay parameter.

Taking the median ISO total polarizability at Camp Beale as our reference, in figure 34 we scale the total polarizability by a factor α ranging between 10^{-5} and 10^{1}, leaving the true value of the decay parameter (decay ≈ 0.18) unaffected. In (a) we show the dependence of the decay distribution on this scaling of target size. As α is decreased, the distribution

becomes increasingly skewed, with the mode of the distribution shifting to smaller values. In (b) and (c) we show that a decrease in target size (smaller α) produces an increased variance in the decay parameter, and an increased positive bias in the expected value. These effects are consistent with what we observe in field data. Smaller targets produce lower SNR data, and so we expect an increase in variance with decreased size. We also find that at later times the signal falls below the noise floor for small targets. This has the effect of biasing the estimate of the decay parameter upwards, as predicted in this analysis.

The top and bottom rows in figure 34 show the effect of varying the correlation coefficient between total polarizabilities at channels t_j and t_k. Increasing the correlation effectively decreases the noise, and reduces the variance and bias of the decay distribution. For the Camp Beale ISOs the correlation coefficient is $R = 1$, so that the expected value of the decay distribution is independent of target size. In figure 34 we reduce the correlation coefficient to $R = 0.99$ and $R = 0.95$. This produces a positive bias in the expected value. When predicting performance for new data sets, the correlation coefficient R estimated from known ISOs will be used.

Figure 35 shows the predicted distributions of the EM61 decay parameter for the same targets considered previously for cued sensor performance prediction (see figure 26). Both targets are uniformly distributed down to 40 cm depth. For this example, the clutter item is slightly slower-decaying than an ISO, and so has a larger expected decay parameter. This is the opposite of the usual situation; clutter tends to be faster decaying). However, we can still predict classification performance - in this case we threshold from smallest to largest decay parameter. The resulting, mediocre ROC is typical of EM-61 performance, with a false alarm rate of 0.51 required to find 99% of the ISOs.

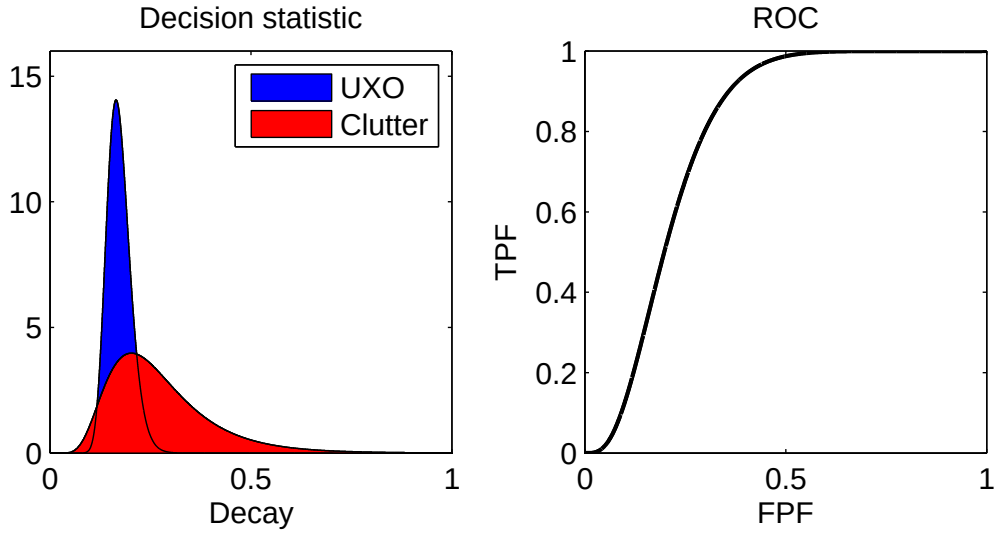


FIGURE 35. Performance prediction for the EM61, Camp Beale noise. Left: distribution of the decay parameter for ISOs and a slow-decaying clutter item (see Camp Beale target 2530, figure 26). Both targets are uniformly distributed down to a maximum depth of 40 cm. Right: predicted receiver operating characteristic.

5. Conclusions and further work

In this report we initially studied the prediction of target response thresholds. Building on previous GSV analyses, we showed that the worst case detection threshold at a specified clearance depth depends on sensor geometry and does not always correspond to a cross-track azimuthal orientation. In addition, we modeled the statistics of anomaly amplitude for a uniform distribution of cross-track location and azimuthal orientation. We found that raising the detection threshold slightly above the worst case scenario can drastically reduce the number of target picks, while still maintaining a high probability of detecting all targets of interest at the maximum clearance depth. Further work will extend this analysis to newer detection sensors such as the OPTEMA and dynamic MetalMapper. We will also examine how the choice of time channels can affect the number of target picks in detection data. We are preparing a journal article on this extended GSV analysis for publication in the next year.

Our work on performance prediction focused on developing realistic models of the noise at each site. We use existing data sets to gain an understanding of how recovered polarizabilities can vary under realistic noise conditions. The physical uniformity of seeded ISO items allows us to isolate the effects of site specific noise on the polarizability distributions. We then use the sample covariance of the ISO polarizabilities to determine the corresponding covariance of the noise for a target at the median estimated location of ISOs in the sample. This noise covariance is highly correlated and does not assume a constant standard deviation for all data at a single channel. With this noise covariance we can then predict the distribution of polarizabilities and ROC that would arise for any specified spatial and angular distributions of targets. Relative to conventional Monte Carlo simulation, this approach to performance prediction is very fast. Indeed, once the forward modeling matrix is pre-computed on a grid of locations, we can instantaneously predict performance for any targets.

To this point we have not tied the estimated covariances of ISO polarizabilities at ESTCP demonstration sites to specific environmental conditions at those sites. Directly linking inputs such as soil conditions and target density to polarizability variance will be a focus of ongoing work. This will allow us to predict performance under arbitrary conditions, rather than restricting ourselves to sites that are, for example, “Beale-like.” Monte Carlo simulations can be used to develop these relationships. For example, we recently simulated the distributions of polarizabilities that will arise in the presence of strongly magnetic Hawaiian soils. These results will be useful for developing an efficient predictive model quantifying the connection between soil response and classification performance.

References

T. Bell, B. Barrow, and J .T. Miller. Subsurface discrimination using electromagnetic induction sensors. *IEEE Trans. Geosci. Remote Sensing*, 39:1286–1293, 2001.

L. S. Beran, S. D. Billings, and D. W. Oldenburg. Robust inversion of time-domain electromagnetic data: application to unexploded ordnance discrimination. *Journal of Engineering and Environmental Geophysics*, 16(3):127–141, 2011.

S. D. Billings, L. R. Pasion, L. Beran, N. Lhomme, L. Song, D. W. Oldenburg, K. Kingdon, D. Sinex, and J. Jacobson. Unexploded ordnance discrimination using magnetic and electromagnetic sensors: Case study from a former military site. *Geophysics*, 75:B103–B114, 2010.

P. Billingsley. *Probability and Measure*. Wiley-Interscience, 1995.

R. M. Brannon. A review of useful theorems involving proper orthogonal matrices referenced to three dimensional physical space, 2002.

J. F. Cardoso. Jacobi angles for simultaneous diagonalization. *SIAM journal on matrix analysis and applications*, 17:161–163, 1996.

L. Carin, L.S. Beran, S.D. Billings, L. R. Pasion, and D. W Oldenburg. Integration of Advanced Statistical Analysis Tools and Geophysical Modeling, SERDP Project MR-1657. Technical report, SERDP, 2012.

T. Hastie, R. Tibshirani, and J. Friedman. *The elements of statistical learning: data mining, inference and prediction*. Spring-Verlag, 2001.

A. A. Kaufman. *Geophysical Field Theory and Method*. Academic Press, 1994.

R. A. Marrona, R. D. Martin, and V. J. Yohai. *Robust Statistics: Theory and Methods*. 2006.

M. S. Munkholm and E. Auken. Electromagnetic noise contamination on transient electromagnetic soundings in culturally disturbed environments. *Journal of Environmental and Engineering Geophysics*, 1(2):119–127, 1996.

H. H. Nelson, K. Kaye, and A. Andrews. Geophysical System Verification (GSV): A Physics-Based Alternative to Geophysical Prove-Outs for Munitions Response. Technical report, ESTCP, 2010.

L. R. Pasion. *Inversion of time-domain electromagnetic data for the detection of unexploded ordnance*. PhD thesis, University of British Columbia, 2007.

L. R. Pasion. Dipole Models for UXO Discrimination at Live Sites Pole Mountain, ESTCP Project MR-201159. Technical report, ESTCP, 2012.

L. R. Pasion, S. D. Billings, D. W. Oldenburg, and S. Walker. Application of a library-based method to time domain electromagnetic data for the identification of unexploded ordnance. *Journal of Applied Geophysics*, 61:279–291, 2007.

W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. *Numerical Recipes in C*. Cambridge University Press, 1992.

M. Prouty, D. C. George, and D. D. Snyder. MetalMapper: A Multi-Sensor TEM System for UXO Detection and Classification, ESTCP Project MR-200603. Technical report, ESTCP, 2011.

F. Shubitidze, B. Barrowes, I. Shamatava, J. P. Fernandez, and K. O'Neill. The orthonormalized volume magnetic source technique applied to live-site UXO data: Inversion and classification studies. *SEG Technical Program Expanded Abstracts*, 30(1):3766–3770, 2011.

L. P. Song, L. R. Pasion, S. D. Billings, and D. W. Oldenburg. Nonlinear inversion for multiple objects in transient electromagnetic induction sensing of unexploded ordnance: Technique and applications. *IEEE Trans. Geosci. Remote Sensing*, 49(10):4007 –4020, 2011.

D. A. Steinhurst, G. R. Harbaugh, J. B. Kingdon, T. Furuya, D. A. Keiswetter, and D. C. George. EMI Array for Cued UXO Discrimination, ESTCP Project MM-0601. Technical report, ESTCP, 2010.

J. R. Wait. *Geo-electromagnetism*. Academic Press, 1982.

G. F. West and J. C. Macnae. *Electromagnetic methods in applied geophysics*, chapter Physics of the electromagnetic exploration method, pages 5–45. SEG, 1991.

D. Williams, Y. Yu, L. Kennedy, X. Zhu, and L. Carin. A bivariate Gaussian model for unexploded ordnance classification with EMI data. *IEEE Geosci. Remote Sensing Letters*, 4:629–633, 2007.

Appendix A. Summary of ESTCP demonstration data sets

Sensors	Targets of Interest	Results
Demonstration Site: Camp Sibert		
• Geonics EM-61 cart • MTADS EM-61 array • MTADS mag array • Cued EM-63	• 4.2" mortars	The results for all sensor combinations were excellent. When inverted cooperatively with magnetics data, the EM63 cued interrogation was the most effective discriminator. All 33 UXO were recovered prior to 25 false alarms. The results from the EM-61 cart were also very good, although 24 false-positives were required to excavate all 105 UXO.
Demonstration Site: San Luis Obispo		
• MTADS mag and EM-61 arrays • Geonics EM61 • MSEMS • TEMTADS • MetalMapper • BUD	• 60 mm mortar • 81 mm mortar • 4.2" mortars • 2.36" rockets • one each of 37 mm, 3" and 5" projectiles	Magnetometer detection and discrimination performance at this site was quite poor. For EM-61 production data, the time-decay rate estimated from the recovered polarizabilities provided an effective ranking scheme. For the TEMTADS data the library method was the most effective with 204 of 206 TOI recovered along with 131 of 1076 non-TOI. The other discrimination methods were also effective, generating between 2 to 4 false-negatives. The library method was again most effective for the MetalMapper data with the excavation of 203 of 204 TOI and 175 of 1205 non-TOI.
Demonstration Site: Camp Butner		
• Geonics EM61 cart • TEMTADS • MetalMapper	• 105 mm projectile • 37 mm projectile • M48 fuze	This site exhibited significant magnetic soil response and relatively high target densities. EM-61 performance was quite poor, with approximately 70 % of non-TOI excavated in order to find all TOI. Excellent classification performance was achieved with TEMTADS data processing, all TOI were readily identified and only 4 % of clutter were dug. The classification method relied on a misfit with respect to all polarizabilities, followed by classification with total polarizability only. MetalMapper performance was comparable to the TEMTADS for identification of 90 % of TOI, but higher noise levels in one of the sensors deployed at this site produced 2-3 missed TOI most performers' diglists.

Sensors	Targets of Interest	Results
Demonstration Site: Pole Mountain		
• Geonics EM61 cart • MetalMapper	• Stokes mortar • 75 mm • 60 mm mortar • 57 mm • 37 mm projectile • Small ISO	This demonstration was divided into two parts (Years 1 and 2). The EM-61 achieved mediocre performance for both parts: the false alarm rates (FAR) for Year 1 and Year 2 were, 64.0 % and 68.0 % respectively. All TOI were recovered at the specified EM-61 operating points. The MetalMapper data produced consistently excellent classification performance. For example, in the Year 1 test the Library method resulted in the excavation of all TOI with a FAR of 2.77 %. A number of analysts processed these data and achieved similar, "near-perfect" performance.
Demonstration Site: Camp Beale		
• Geonics EM61 cart • MetalMapper • MPV • TEMTADS2x2 • BUDHH	• 105 mm • 81 mm mortar • 60 mm mortar • 37 mm projectile • Small ISO	This demonstration tested classification in a portion of the site amenable to vehicular towed systems (i.e. the "Open Area"). In addition, man-portable sensor data were collected with the TEMTADS 2x2, BUD, and MPV sensors in a treed section of the site (i.e. the "Portable Area"). For the Portable Area, the EM-61 analysis required approximately 30 % of clutter to be excavated in order to find all TOI. For the Open Area, the false alarm rate was 50 % with this sensor. MetalMapper data were acquired by Parsons and CH2M HILL in the Open Area, with a different analyst processing each data set. Application of a library classification method to the CH2M HILL data resulted in 99.2 % of TOI identified (due to one missed TOI) and 27.4 % scrap dug at the stop dig point. For the Beale Parsons data two stage (all polarizabilities/total polarizability) classification with a support vector machine was able to identify all TOI at the stop-dig point, with a 22 % false alarm rate. All portable sensors data sets produced diglists with no TOI left in the ground in the portable area of the site. False alarm rates were also quite similar for all portable sensors, averaging approximately 25 % of clutter dug in order to find all TOI.

Appendix B. Expected polarizability covariance under a uniform distribution of target orientations

To predict classification performance in section 4 we use a single target orientation (horizontal or vertical) to propagate uncertainties and predict the distribution of the decision statistic. More realistically, we can model a specified distribution of target orientations. A uniform distribution of target azimuth and dip is the obvious choice, and here we develop analytic expressions for the covariance of the estimated polarizabilities under a uniform rotation.

The symmetric polarizability tensor $\mathbf{M}$ is related to the diagonal matrix of principal polarizabilities $\mathbf{L}$ via the Euler rotation matrix $\mathbf{A}$

$$\mathbf{M}(t) = \mathbf{A}\mathbf{L}(t)\mathbf{A}^T. \tag{45}$$

For target dip $b \in [0\ \pi]$ and azimuth $g \in [0\ 2\pi]$, we define $\mathbf{A}$ as

$$\mathbf{A} = \begin{pmatrix} \cos(g)\cos(b) & \sin(g) & -\cos(g)\sin(b) \\ -\sin(g)\cos(b) & \cos(g) & \sin(g)\sin(b) \\ \sin(b) & 0 & \cos(b) \end{pmatrix}. \tag{46}$$

Under this convention the diagonal polarizability matrix $\mathbf{L}$ is ordered

$$\mathbf{L}(t) = \begin{pmatrix} L_2(t) & 0 & 0 \\ 0 & L_3(t) & 0 \\ 0 & 0 & L_1(t) \end{pmatrix} \tag{47}$$

so that a vertical target with $g = 0$ has its primary polarizability ($L_1(t)$) aligned with the z axis. The predicted data $\mathbf{d}$ for a target at location $\mathbf{r}$ are given by

$$\mathbf{d} = \mathbf{G}(\mathbf{r})\mathbf{m} \tag{48}$$

with the model vector $\mathbf{m}$ comprised of the unique elements of the polarizability tensor

$$\mathbf{m} = \begin{pmatrix} M_{11} \\ M_{12} \\ M_{13} \\ M_{22} \\ M_{23} \\ M_{33} \end{pmatrix} = \begin{pmatrix} L_2\cos(b)^2\cos(g)^2 + L_1\cos(g)^2\sin(b)^2 + L_3\sin(g)^2 \\ -L_2\cos(g)\sin(g)\cos(b)^2 - L_1\cos(g)\sin(g)\sin(b)^2 + L_3\cos(g)\sin(g) \\ L_2\cos(b)\cos(g)\sin(b) - L_1\cos(b)\cos(g)\sin(b) \\ L_2\cos(b)^2\sin(g)^2 + L_3\cos(g)^2 + L_1\sin(b)^2\sin(g)^2 \\ L_1\cos(b)\sin(b)\sin(g) - L_2\cos(b)\sin(b)\sin(g) \\ L_1\cos(b)^2 + L_2\sin(b)^2 \end{pmatrix} \tag{49}$$

with the dependence on time t now implicit. In this formulation the forward modeling matrix $\mathbf{G}$ has dimensions $N \times 6$, with N the number of data measured at a single time channel.

From the above expressions, we define the operator α

$$\alpha = \begin{pmatrix} \cos(b)^2\cos(g)^2 & \sin(g)^2 & \cos(g)^2\sin(b)^2 \\ -\cos(b)^2\cos(g)\sin(g) & \cos(g)\sin(g) & -\cos(g)\sin(b)^2\sin(g) \\ \cos(b)\cos(g)\sin(b) & 0 & -\cos(b)\cos(g)\sin(b) \\ \cos(b)^2\sin(g)^2 & \cos(g)^2 & \sin(b)^2\sin(g)^2 \\ -\cos(b)\sin(b)\sin(g) & 0 & \cos(b)\sin(b)\sin(g) \\ \sin(b)^2 & 0 & \cos(b)^2 \end{pmatrix} \tag{50}$$

such that α maps from the principal polarizabilities to unique elements of the polarizability tensor

$$\mathbf{m} = \alpha \begin{pmatrix} L_2 \\ L_3 \\ L_1 \end{pmatrix}. \tag{51}$$

Now we assume that the covariance of the data for a given target orientation is given by a transformation of the polarizability covariance

$$\mathrm{cov}(\mathbf{d}) = \mathbf{G}_{eq}\ \alpha\ \mathrm{cov}(\mathbf{L})\ \alpha^T\ \mathbf{G}_{eq}^T \tag{52}$$

with $\mathbf{G}_{eq}$ corresponding to the median location (relative to the sensor) of all targets used to compute $\mathrm{cov}(\mathbf{L})$. Now for a target at location $\mathbf{r}$ the covariance of the estimated polarizabilities can be expressed as

$$\begin{aligned} \mathrm{cov}(\hat{\mathbf{L}}) =& \beta\ \mathbf{G}^\dagger(\mathbf{r})\ \mathrm{cov}(\mathbf{d})\ (\mathbf{G}^\dagger(\mathbf{r}))^T\ \beta^T \\ =& \beta\ \mathbf{G}^\dagger(\mathbf{r})\ \mathbf{G}_{eq}\ \alpha\ \mathrm{cov}(\mathbf{L})\ \alpha^T\ \mathbf{G}_{eq}^T\ (\mathbf{G}^\dagger(\mathbf{r}))^T\ \beta^T \end{aligned} \tag{53}$$

where the pseudo-inverse is $\mathbf{G}^\dagger = (\mathbf{G}^T\mathbf{G})^{-1}\mathbf{G}$. The operator β

$$\beta = \begin{pmatrix} \cos(b)^2\cos(g)^2 & \sin(g)^2 & \cos(g)^2\sin(b)^2 \\ -2\cos(b)^2\cos(g)\sin(g) & 2\cos(g)\sin(g) & -2\cos(g)\sin(b)^2\sin(g) \\ \cos(b)^2\sin(g)^2 & 0 & -2\cos(b)\cos(g)\sin(b) \\ 2\cos(b)\cos(g)\sin(b) & \cos(g)^2 & \sin(b)^2\sin(g)^2 \\ 2\cos(b)\sin(b)\sin(g) & 0 & 2\cos(b)\sin(b)\sin(g) \\ \sin(b)^2 & 0 & \cos(b)^2 \end{pmatrix}^T \tag{54}$$

corresponds to diagonalization of the estimated polarizability tensor

$$\hat{\mathbf{L}} = \mathbf{A}^T\hat{\mathbf{M}}\mathbf{A}, \tag{55}$$

so that

$$\begin{pmatrix} \hat{L}_2 \\ \hat{L}_3 \\ \hat{L}_1 \end{pmatrix} = \beta\hat{\mathbf{m}}. \tag{56}$$

In addition, since $\mathbf{A}$ is orthogonal, we find that

$$\beta\alpha = \mathbf{I}. \tag{57}$$

Equation 53 provides an expression for the covariance of the estimated polarizabilities at a given orientation. We now assume a uniform probablility distribution of target dip b and azimuth g, with the independent densities (Brannon, 2002)

$$\begin{aligned} p(g) &= \frac{1}{2\pi} \\ p(b) &= \frac{1}{2}\sin(b). \end{aligned} \tag{58}$$

We marginalize equation 53 over azimuth and dip to obtain the expected covariance of the polarizbility estimates

$$E(\mathrm{cov}(\hat{\mathbf{L}})) = \frac{1}{4\pi}\int_{g=0}^{2\pi}\int_{b=0}^{\pi} \mathrm{cov}(\hat{\mathbf{L}})\sin(b)\ db\ dg \tag{59}$$

These integrals involve only trigonometric functions and can be evaluated analytically. The resulting expressions are quite lengthy and so we do not reproduce them here.