FINAL REPORT

Analysis of Next Generation Sensor Data

SERDP Project MR-1662

OCTOBER 2021

Dean Keiswetter
Leidos

Barry Spargo
Brandy White
Naval Research Laboratory

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*
31/10/2021

2. REPORT TYPE
SERDP Final Report

3. DATES COVERED *(From - To)*
6/15/2009 - 8/30/2013

4. TITLE AND SUBTITLE
Analysis of Next Generation Sensor Data

5a. CONTRACT NUMBER
09-C-0011

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Leidos
Dean Keiswetter

Naval Research Laboratory
Barry Spargo
Brandy White

5d. PROJECT NUMBER
MR-1662

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Leidos
120 Quade Dr.
Cary, NC 27513

8. PERFORMING ORGANIZATION REPORT NUMBER
MR-1662

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Strategic Environmental Research and Development Program (SERDP)
4800 Mark Center Drive, Suite 16F16
Alexandria, VA 22350-3605

10. SPONSOR/MONITOR'S ACRONYM(S)
SERDP

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
MR-1662

12. DISTRIBUTION/AVAILABILITY STATEMENT
DISTRIBUTION STATEMENT A. Approved for public release: distribution unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

An algorithm was developed to enumerate, locate, and characterize individual signal sources given observation of their combined signals. No a-priori estimate for the number of sources is required. We assume a forward model exists, and that superposition holds, i.e. coupling between sources is ignored. A system of linear equations $y = Ax$ is set up in which columns of matrix A contain expected signals from many hypothesized sources, and y contains the observed signal. Recently developed solvers designed for linear systems with sparse non-negative solutions make this approach feasible even when large numbers of sources are involved. With each iteration, the collection of hypothesized sources is refined using a Harmony Search algorithm. Application is demonstrated on the problem of locating multiple buried conductors based on electromagnetic induction (EMI) signals observed at ground surface.

15. SUBJECT TERMS
Analysis, Next Generation Sensor Data

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Dean Keiswetter
UNCLASS	UNCLASS	UNCLASS	UNCLASS	22	**19b. TELEPHONE NUMBER** *(Include area code)* 919-677-1560

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18

FINAL REPORT

Project: MR-1662

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

ABSTRACT

An algorithm was developed to enumerate, locate, and characterize individual signal sources given observation of their combined signals. No a-priori estimate for the number of sources is required. We assume a forward model exists, and that superposition holds, i.e. coupling between sources is ignored. A system of linear equations $y = Ax$ is set up in which columns of matrix A contain expected signals from many hypothesized sources, and y contains the observed signal. Recently developed solvers designed for linear systems with sparse non-negative solutions make this approach feasible even when large numbers of sources are involved. With each iteration, the collection of hypothesized sources is refined using a Harmony Search algorithm. Application is demonstrated on the problem of locating multiple buried conductors based on electromagnetic induction (EMI) signals observed at ground surface.

EXECUTIVE SUMMARY

Introduction

As a result of past military training and weapons-testing activities, military munitions, including unexploded ordnance, are present at sites designated for base realignment and closure and at Formerly Used Defense Sties. Efforts to build sensors and analysis systems that efficiently remediate military munitions are beginning to show promise. Efficiency is achieved through classification of military munitions from the multitude of non-hazardous metallic fragments and clutter that are often located very near and around munitions items. Subsurface military munitions are rare compared to non-hazardous munitions debris.

Classification of military munitions relies on the detailed analysis of broadband electromagnetic induction (EMI) data. This analysis becomes more complicated if multiple metallic sources are unknowingly within the EMI sensors footprint.

Objective

The objective of this project was to develop an algorithm that automatically determines the number of sources and their respective locations based on measure EMI sensor data.

Technical Approach

This algorithm is designed to locate and identify buried conductive metal targets based on EMI signals collected at ground surface, and it addresses the case of multiple targets spaced so closely together that their signals overlap, and it is not possible to get isolated solutions on individual targets. The algorithm proceeds by successively refining a large collection of hypothesized point sources, whose number (parameter p) is typically 100 or more, intended to be larger than the number of actual buried targets present. As the algorithm iterates, the spatial distribution of sources evolves to produce better and better agreement with the observed EMI signals. After convergence, a clustering algorithm is run on the cloud of sources, and prominent clusters are identified as individual buried targets.

This algorithm relies on the dipole model, which assumes each source produces a transient dipolar secondary field proportional to the transmitted primary field, where the coefficient of proportionality is a symmetric 3x3 tensor with 6 unique elements. Each source is therefore fully defined by nine values: three position coordinates and six tensor entries. Interaction between sources is ignored i.e. source responses are calculated purely from the primary field, not from any secondary fields arising from other sources nearby.

The locations and tensor entries of the sources are all allowed to vary, which means the model has 9 x p (i.e. many hundreds) of fitted parameters, which means it's extremely flexible and non-physical solutions are easy to obtain, so regularization is needed, whereby external information is used to limit the way parameters may be adjusted. This is done using the QHull tetrahedron procedure.

In each iteration, the dipole model is run on each source to calculate a 1-D signal vector of length M for each source, representing expected signals for that source. For example, the NRLTEM array had 25 receive coils to pick up the secondary fields that resulted from 25 separate transmit coils fired sequentially, so the 1-D signal vector had length M=625 (25 x 25) representing the expected signal for an individual source.

The NRL TEM array also records the time-decay of the secondary field over several timegates, which makes the signal vector M even bigger, but this algorithm is aiming to locate thespatial positions of the buried targets, and the observed EMI signal captures the same spatial information of each time-gate, since buried targets don't move. In experience, we found that performance does not improve when the signal vector includes multiple time gates, and target locations are recovered best by simply summing time-gate information over a fixed interval of the decay curve, to create the observed EMI signal used to drive evolution of the source cloud.

Signal vectors for each source are calculated and assembled into a 2-D array "A" with M rows and p columns, where each column represents the expected signal from an individual source. Next, we find x, a vector of weights (length p) which are applied to the sources and produce best-fit match to the observed EMI signal, i.e. we solve $y_{obs}{\sim}y=Ax$, where y is the summed signal from all the weighted sources which best-fits the true observed signal y_{obs}. We require that the weights x must be non-negative (no such thing as a negative buried object) and also, we want a parsimonious solution i.e. most weights should be zero.

At the end of each iteration, we have a collection of p hypothesized sources, each having its own position and response tensor, along with non-negative weight vector x, which gives the best match the observed EMI signal. To refine the sources for the next iteration, we want to condense sources together where possible, to create a smaller number of new sources which closely match the same EMI signal produced by the larger number of old sources that received weight in the previous iteration. For example, if the solver gives weight to a few closelyspaced sources, this suggests there may be a true target lying within the space they encompass. We aim to condense these weighted sources into a single new source which may better approximate the true target, in the hopes that this new source will receive weight in the next iteration and provide an improved estimate of this target position.

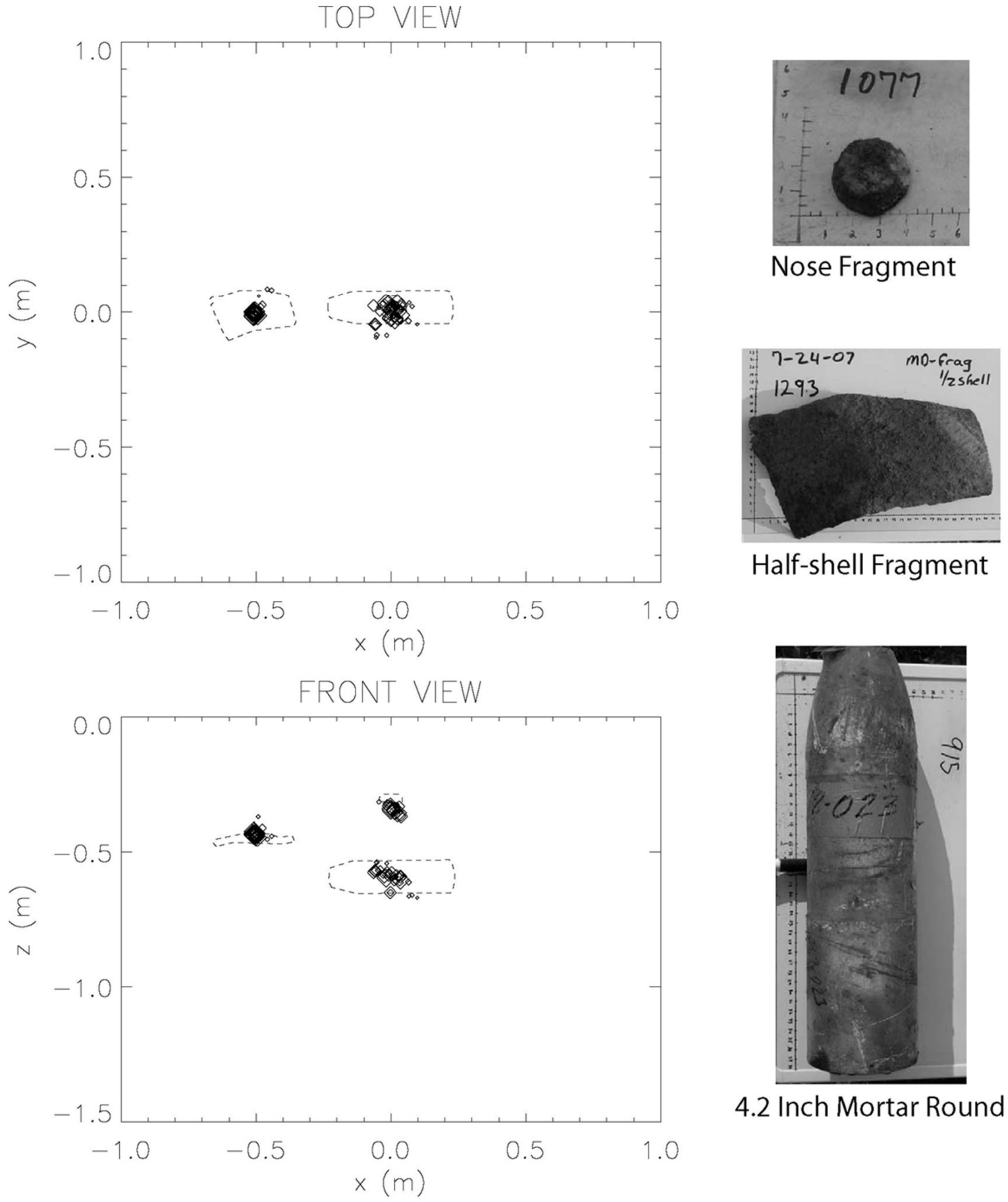
To accomplish this condensation, we take the spatial positions of all the sources which received weights, and submit them all to the QHull algorithm, which partitions the space into tetrahedrons. We cycle through each tetrahedron and 1) find the centroid, which is simply the weighted average of the four vertices, then 2) create a new source at the centroid and adjust its tensor entries so that it best-matches the combined response of the four vertices. This is the only way that tensor entries are permitted to be adjusted, and it comprises the regularization described above.

Results and Discussion

These data were collected with the Naval Research Lab (NRL) Time-domain Electro-Magnetic induction (TEM) array, a state-of-the-art sensor with 25 transmitter coils and 25 receiver coils, each square in shape and about 40cm on a side, arranged in a 5 by 5 grid on a horizontal plane, making the overall instrument about 2 by 2 meters in size. This system is designed to discriminate buried unexploded ordnance (UXO) from harmless clutter, to reduce the cost of cleanup at former Department of Defense sites.

The TEM operates by pulsing current sequentially through the 25 transmitter coils to develop transient magnetic fields in the sub-surface, which induce eddy currents in buried conductors, giving rise to secondary magnetic fields that are detected in the receive coils. Analysis of these signals is based on the dipole model, which assumes each buried conductor behaves as a point source, producing a transient dipolar field proportional to the transmitted primary field, where the coefficient of proportionality is a symmetric tensor with 6 unique elements. Each source is therefore fully characterized by nine values: three position coordinates and six tensor entries.

Three data sets are analyzed, representing three different kinds of data: the first is test-stand-data collected indoors at Blossom Point, MD, with the TEM array held approximately 6 feet off the ground on non-conducting supports, and metal targets carefully positioned below. This arrangement allows for testing the instrument without soil effects, and with accurate knowledge of the true distribution of sources. The second data set is field data from Aberdeen Proving ground MD (APG). It was collected on calibration target A1, which is a 155mm artillery round buried at about 1 m depth. Here, soil is present, and the true target location is less accurately known. The third data set is the same A1 target from APG, but with four additional synthetic clutter targets added into the data. Figures 1 through 3 illustrate results on these data.

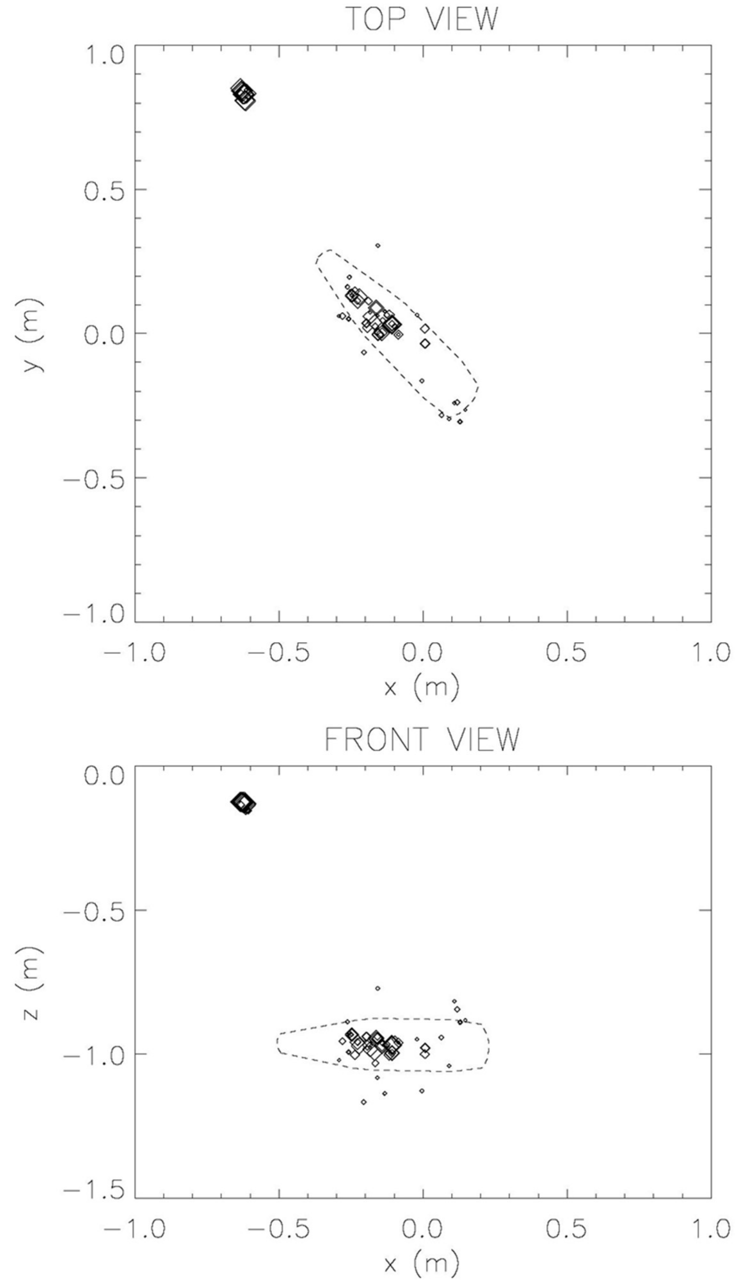


Nose Fragment

Half-shell Fragment

4.2 Inch Mortar Round

Figure 1. Results on data from Blossom Point collected in 2007 with the NRL TEM array. This test was conducted in air above ground with the sensor raised on non-conducting supports. The lowest object is a horizontal 4.2 inch mortar round, the next is a flattened “half-round” fragment, and the shallowest is a nose fragment. Black diamonds represent weights assigned to hypothetical sources by the algorithm, and the size of the diamond is proportional to the magnitude of the weight. Only the heaviest quartile of weights are shown. Dashed lines indicate true target locations. These results were obtained with the number of sources *p* set to 100, the size of the Harmony memory *k* set to 20, and the maximum number of iterations *nmax* set to 100. Processing time was under 20 seconds on a Dell PWS690 workstation.



This 155mm artillery round is similar to the one detected below ground at Aberdeen Proving Ground MD.

Figure 2. Results on data from Aberdeen Proving Ground (APG) collected in 2008 with the NRL TEM array. This is target A1 in the calibration grid at APG, which is a horizontal 155mm projectile centered approximately 0.7 meters below ground surface (z axis in these figures shows distance from the TEM sensor, not ground surface). As before, black diamonds represent weights assigned by the algorithm – heaviest quartile only - and the dashed line indicates the known target. The diamonds in the upper left corner are caused by an unknown shallow clutter item, apparently resting on ground surface. This clutter object complicates analysis of these data using standard single-target inversion methods to the point where the buried ordnance cannot be identified. These results were obtained in under 20 seconds on a Dell PWS690 workstation with parameters p =100*,* k=20, and $nmax$=100*,* as in Figure 1.

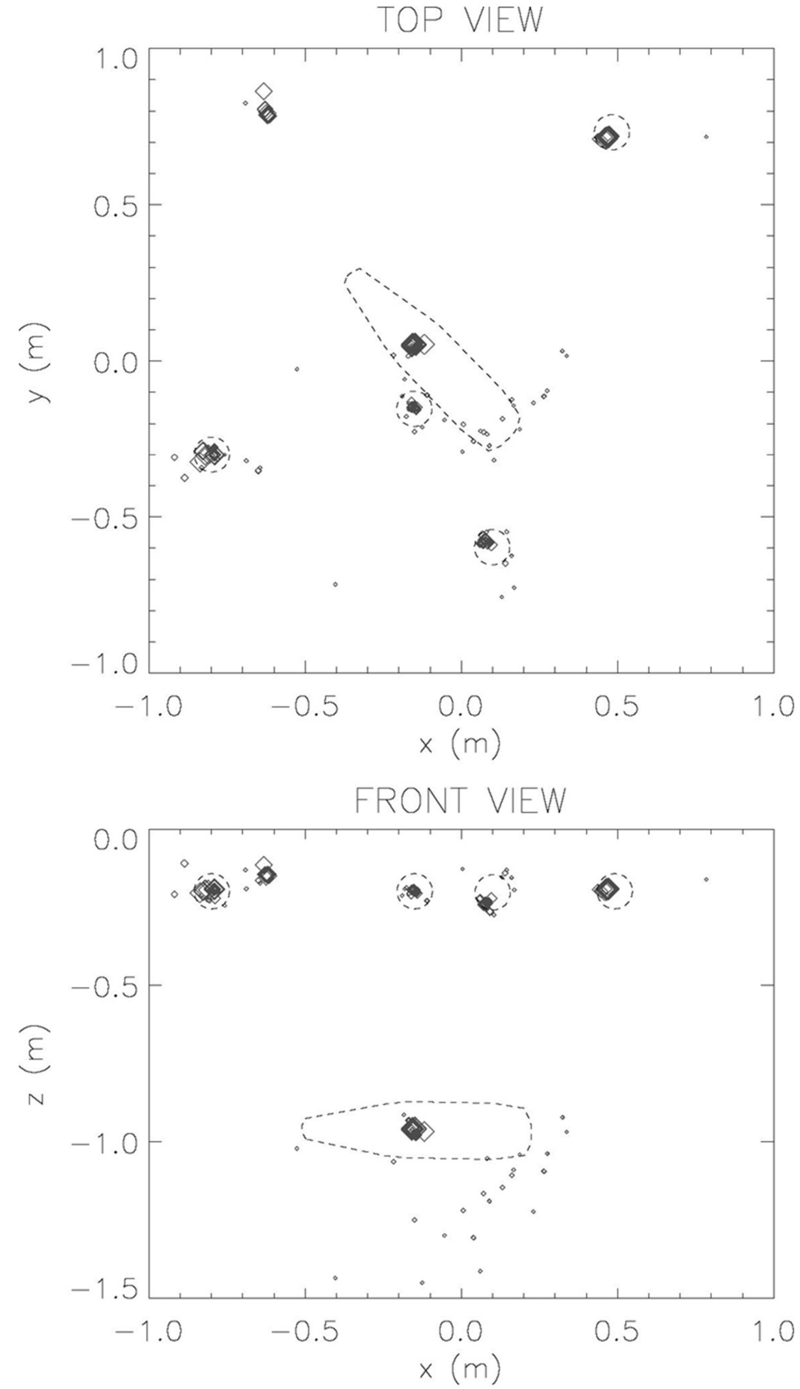


Figure 3. Four shallow synthetic sources were added into the APG data of figure 2. These four sources are spheres, all resting 20cm below the sensor, intended to represent near-surface clutter typically seen at active cleanup sites. The algorithm correctly located all six targets in this example, including the deep 155mm artillery round and the unknown shallow clutter in the upper left. Algorithm parameters were adjusted upwards somewhat in this example to allow for the more complicated environment: these were *p* =200*,* *k*=20, and *nmax*=400. Processing time was under 40 seconds. Again, only the heaviest quartile of weights are shown (black diamonds), and dashed lines represent known sources.

Implications for Future Research and Benefits

An algorithm has been presented and demonstrated to locate and characterize multiple signal sources based on observation of their combined signals. The method relies on new solvers designed to provide sparse non-negative solutions for underdetermined linear systems of equations, permitting many hypothetical sources to be evaluated jointly with each iteration. Most of these sources receive zero weight, but those with positive weight drive successive refinement of the collection. A crucial consideration is the grouping of non-negative weights into local neighborhoods to estimate true sources which may be located inside. This grouping is done by partitioning the volume into tetrahedrons, each vertex of which is a non-zero source that received weight in the preceding iteration. Approximation of the presumed true source inside is accomplished by locating the tetrahedron centroid and then assigning source parameters to best match the summed signal from the four vertices. The overall number of sources in the system is indicated by the number of clusters in the output, which may be detected by clustering algorithms capable of counting clusters, such as Quality Threshold (QT) clustering.

Application to the problem of electromagnetic induction is demonstrated, using signals from the NRL TEM array. Results show successful location of multiple sources for both in-air test stand measurements and field measurements.

Future research is needed in speed improvements, automatic background removal, and the application to data acquired while moving.

OBJECTIVE

The objective of this project was to develop an algorithm that automatically determines the number of sources and their respective locations based on measure EMI sensor data.

BACKGROUND

This algorithm is designed to locate and identify buried conductive metal targets based on EMI signals collected at ground surface, and it addresses the case of multiple targets spaced so closely together that their signals overlap, and it is not possible to get isolated solutions on individual targets. The algorithm proceeds by successively refining a large collection of hypothesized point sources, whose number (parameter p) is typically 100 or more, intended to be larger than the number of actual buried targets present. As the algorithm iterates, the spatial distribution of sources evolves to produce better and better agreement with the observed EMI signals. After convergence, a clustering algorithm is run on the cloud of sources, and prominent clusters are identified as individual buried targets.

This algorithm relies on the dipole model, which assumes each source produces a transient dipolar secondary field proportional to the transmitted primary field, where the coefficient of proportionality is a symmetric 3x3 tensor with 6 unique elements. Each source is therefore fully defined by nine values: three position coordinates and six tensor entries. Interaction between sources is ignored i.e. source responses are calculated purely from the primary field, not from any secondary fields arising from other sources nearby.

The locations and tensor entries of the sources are all allowed to vary, which means the model has 9 x p (i.e. many hundreds) of fitted parameters, which means it's extremely flexible and non-physical solutions are easy to obtain, so regularization is needed, whereby external information is used to limit the way parameters may be adjusted. This is done using the QHull tetrahedron procedure.

In each iteration, the dipole model is run on each source to calculate a 1-D signal vector of length M for each source, representing expected signals for that source. For example, the NRLTEM array had 25 receive coils to pick up the secondary fields that resulted from 25 separate transmit coils fired sequentially, so the 1-D signal vector had length M=625 (25 x 25) representing the expected signal for an individual source.

The NRL TEM array also records the time-decay of the secondary field over several timegates, which makes the signal vector M even bigger, but this algorithm is aiming to locate thespatial positions of the buried targets, and the observed EMI signal captures the same spatial information of each time-gate, since buried targets don't move. In experience, we found that performance does not improve when the signal vector includes multiple time gates, and target locations are recovered best by simply summing time-gate information over a fixed interval of the decay curve, to create the observed EMI signal used to drive evolution of the source cloud.

Signal vectors for each source are calculated and assembled into a 2-D array "A" with M rows and p columns, where each column represents the expected signal from an individual source. Next, we find x, a vector of weights (length p) which are applied to the sources and produce best-fit match to the observed EMI signal, i.e. we solve $y_{obs} \sim y=Ax$, where y is the summed signal from all the weighted sources which best-fits the true observed signal y_{obs}. We require that the weights *x* must be non-negative (no such thing as a negative buried object) and also, we want a parsimonious solution i.e. most weights should be zero.

At the end of each iteration, we have a collection of *p* hypothesized sources, each having its own position and response tensor, along with non-negative weight vector *x*, which gives the best match the observed EMI signal. To refine the sources for the next iteration, we want to condense sources together where possible, to create a smaller number of new sources which closely match the same EMI signal produced by the larger number of old sources that received weight in the previous iteration. For example, if the solver gives weight to a few closelyspaced sources, this suggests there may be a true target lying within the space they encompass. We aim to condense these weighted sources into a single new source which may better approximate the true target, in the hopes that this new source will receive weight in the next iteration and provide an improved estimate of this target position.

To accomplish this condensation, we take the spatial positions of all the sources which received weights, and submit them all to the QHull algorithm, which partitions the space into tetrahedrons. We cycle through each tetrahedron and 1) find the centroid, which is simply the weighted average of the four vertices, then 2) create a new source at the centroid and adjust its tensor entries so that it best-matches the combined response of the four vertices. This is the only way that tensor entries are permitted to be adjusted, and it comprises the regularization described above.

MATERIALS AND METHODS

We assume an individual source *i* is fully described by an *N*-dimensional feature vector $X_i \in \Re^N$ which includes spatial coordinates, and the associated *M*-dimensional signal vector $S_i \in \Re^M$ is generated through some sensor system. Both X and S are assumed to be continuous and the mapping $X \Rightarrow S$ is assumed to be smooth. Under these conditions, and given that superposition holds, the feature vector of an individual source X_i can be approximated as a weighted average of similar sources in a small neighborhood:

$$X_i \cong \sum_j w_j X_j \qquad j \in \text{neighborhood around } X_i, \tag{1}$$

and the corresponding signal S_i may be approximated using the same weights applied to signals from that neighborhood:

$$S_i \cong \sum_j w_j S_j \qquad j \in \text{neighborhood around } S_i. \tag{2}$$

This suggests a possible scheme for estimating X_i through successive approximation:

1. Generate many hypothetical sources X_j near an estimated source of interest $X_{ESTIMATE}$.
2. Run the forward model to find the expected signal S_j for each X_j.
3. Find the weights w_j to best approximate the observed signal: $S_{OBSERVED} \cong \sum_j w_j S_j$

4. Use those weights to update the estimate: $X_{NEW\ ESTIMATE} = \sum_j w_j X_j$.

A difficulty in this scheme is step 3. The weights must be non-negative, which rules out efficient solvers like SVD, and to facilitate convergence and avoid local minima one needs to include a large number of hypothetical sources - preferably many more than the length of the signal vector, but then the system becomes underdetermined and multiple solutions are possible. Techniques for solving underdetermined linear systems of equations have recently focused on the phenomenon of *L0/L1 equivalence*, whereby the L0 solution– the one having fewest nonzero, is also likely to be the L1 solution – the one with the smallest magnitude, provided the solution vector is sparse enough, which in practice means no more than about 30% nonzero for many problems [Donoho, Donaho and Tanner, 2005]. This is important since the L1 solution is relatively easy to obtain, while L0 is difficult, and numerous applications, driven by parsimony, require the L0 solution. Thanks to L0/L1 equivalence, we can get the *one* solution, among all possible, which has the fewest non-zeros, and this facilitates convergence of the successive approximation scheme outlined above.

The algorithm proceeds iteratively by updating a large collection of hypothesized sources, assumed to be greater in number than the actual number present to ensure sparse solutions. With each iteration, signals from all the hypothesized sources are offered to the linear solver, and only a small fraction receive weights. The algorithm then updates the collection based on those assigned weights to home-in on the true distribution of sources.

Partitioning Feature Space

When multiple sources are in play, the weights w in step 4 obviously cannot all be used at once to update a single source, but instead must be grouped into local neighborhoods each encompassing a presumed true source located somewhere within. This is done by partitioning N-dimensional feature space into polytopes with vertices defined by the feature vectors X_j that received weights from the solver in step 3. For example, if the feature vector X contains only spatial coordinates (N=3), then the polytopes are tetrahedrons and the running index j takes just 4 values (N+1) as it cycles through the vertices to estimate the location of the presumed source inside. Each tetrahedron would generate a separate estimated source in this way, and all of these would be offered to the solver in the pending iteration.

A Practical Short Cut

Description of a source typically requires more than just the three spatial coordinates. In the electromagnetic induction case given below, it takes nine. In practice we have not found it feasible to partition a feature space with more than three or four dimensions, so an expedient "short cut" is employed whereby only the three spatial coordinates of the feature vector X are used. The polytopes are therefore tetrahedrons as described above, which permits much faster processing since the QHull algorithm can be used, which is very efficient [Barber et all, 1996]. Also, only four vertices X_j need be considered when assigning parameters to the newly estimated source at the centroid, as summarized in step 4 above. A potential disadvantage is poor accuracy. Since grouping is based solely on spatial position, it is conceivable that when true sources are closely spaced, vertices may be grouped into the wrong local neighborhoods and therefore confuse the updating process, but in practice this potential problem is apparently minimal. With this shortcut, assignment of spatial position for the new source is straightforward, and the other elements of the feature vector X are determined using a quick parameter search such that the signal from the new source approximates the weighted sum of signals from the vertices. To summarize, spatial coordinates for the newly estimated source are determined by the tetrahedron centroid, and any other features of the source are assigned as follows:

Sum the weighted signals from the four vertices to find the 'goal' signal:

$$S_{GOAL} = \sum_j w_j S_j, \quad j = 1..4.$$

Assign remaining features of the new source to best match the 'goal':

$$X_{NEW} \Rightarrow S_{NEW} \cong S_{GOAL}$$

In the electromagnetic induction example given later, the six unique tensor elements which, combined with the three spatial coordinates fully define the source, are assigned so as to best match the sum of weighted signals from the four vertices of the encompassing tetrahedron. This process is repeated for each tetrahedron in the partition and all the new sources produced in this way are stored and offered to the solver in the pending iteration.

Harmony Search

The Harmony Search algorithm is so named after the improvisation process of musicians, in which each plays a single note aimed at finding best overall harmony with the ensemble [Kang and Geem, 2005]. This is analogous to each decision variable in an optimization problem taking a single value aimed at finding a best overall global optimum of some objective function, and in our case the decision variables correspond to individual sources in the multi-source environment and the objective function is the overall model match to observed data. In the standard Harmony Search algorithm, a running record of previous trials is maintained, termed the "harmony memory", and this record forms the basis for generating new trials. Values for each decision variable in a new trial are assigned one-by-one using three possible methods: (a) assign a previously successful value retrieved from harmony memory, (b) use a small variation around a previously successful value, or (c) use an entirely random value. The choice (a) through (c) is determined randomly for each decision variable and each new trial. Once assembled, the new trial is tested by the objective function, and if it performs better than the worst performer in the harmony memory, then it replaces that entry and becomes a possible basis for future trials. The current multi-target inversion algorithm uses some of the main features of Harmony Search, namely the running record, variations around previously successful values, and random assignments, but each new trial is based only on a single previous trial from the running record – not a combination of trials. Testing has revealed this to be more effective than standard Harmony Search.

Algorithm Summary

The proposed algorithm therefore has the following steps:

- *Initialize the harmony memory*: generate k trials, each consisting of p sources, and evaluate for each the objective function, based on goodness of fit with the observed signal. Evaluation includes assignment of weights to the sources, most of which will be zero.
- *Select a trial and partition the source space:* select one entry from harmony memory and use the locations of the sources – the ones that received non-zero weight - as vertices of tetrahedrons to partition the domain spatially.

- *Determine centroids:* for each tetrahedron, find the weighted average of the vertex locations. This is the location estimate for the potential true source within.
- *Determine best-fit parameters for centroid:* For each tetrahedron, find the expected combined signal from the four vertices and adjust the remaining parameters of the centroid source (beyond the x,y,z coordinate parameters) to best-fit that signal.
- *Create new trial;* cycle through all the tetrahedrons and gather the new centroid sources into a new trial. Append additional random sources as needed until the desired number p is met.
- *Evaluate the trial*: Invoke the linear solver to find the best match with the observed signal. If the match is better than the worst in harmony memory, swap in this trial.
- *Repeat* 2 thru 6 until a convergence criterion has been met or the maximum number of iterations *nmax* has been reached.

The parameters of this algorithm are:

- k, the number of trials kept in the harmony memory. Typically, 1 to 50.
- p, the number of sources in each trial. Typically, 100 or more.
- *nmax*, the maximum number of iterations. Typically, a few hundred, up to a few thousand.

The selection that occurs is done simply by cycling through the harmony memory sequentially. This ensures that no entry is missed.

Enumerating sources

Output takes the form of spatially distributed weights, and the number and location of sources in the system is indicated by the number and location of clusters in the distribution. Clusters are clearly evident when the signal is dominated by individual isolated sources, but may also be ambiguous if for example the true sources are arranged in a large diffuse cloud. The number and location of sources is recovered by processing output through a clustering algorithm such as Quality Threshold (QT) clustering which requires no prior information regarding the number of clusters present. Results from the following demonstration clearly show the correspondence between clusters and sources.

RESULTS AND DISCUSSION

These data were collected with the Naval Research Lab (NRL) Time-domain Electro-Magnetic induction (TEM) array, a state-of-the-art sensor with 25 transmitter coils and 25 receiver coils, each square in shape and about 40cm on a side, arranged in a 5 by 5 grid on a horizontal plane, making the overall instrument about 2 by 2 meters in size. This system is designed to discriminate buried unexploded ordnance (UXO) from harmless clutter, to reduce the cost of cleanup at former Department of Defense sites.

The TEM operates by pulsing current sequentially through the 25 transmitter coils to develop transient magnetic fields in the sub-surface, which induce eddy currents in buried conductors, giving rise to secondary magnetic fields that are detected in the receive coils. Analysis of these signals is based on the dipole model, which assumes each buried conductor behaves as a point source, producing a transient dipolar field proportional to the transmitted primary field, where the coefficient of proportionality is a symmetric tensor with 6 unique elements. Each source is therefore fully characterized by nine values: three position coordinates and six tensor entries.

Three data sets are analyzed, representing three different kinds of data: the first is test-stand-data collected indoors at Blossom Point, MD, with the TEM array held approximately 6 feet off the ground on non-conducting supports, and metal targets carefully positioned below. This arrangement allows for testing the instrument without soil effects, and with accurate knowledge of the true distribution of sources. The second data set is field data from Aberdeen Proving ground MD (APG). It was collected on calibration target A1, which is a 155mm artillery round buried at about 1 m depth. Here, soil is present, and the true target location is less accurately known. The third data set is the same A1 target from APG, but with four additional synthetic clutter targets added into the data. Figures 1 through 3 illustrate results on these data.

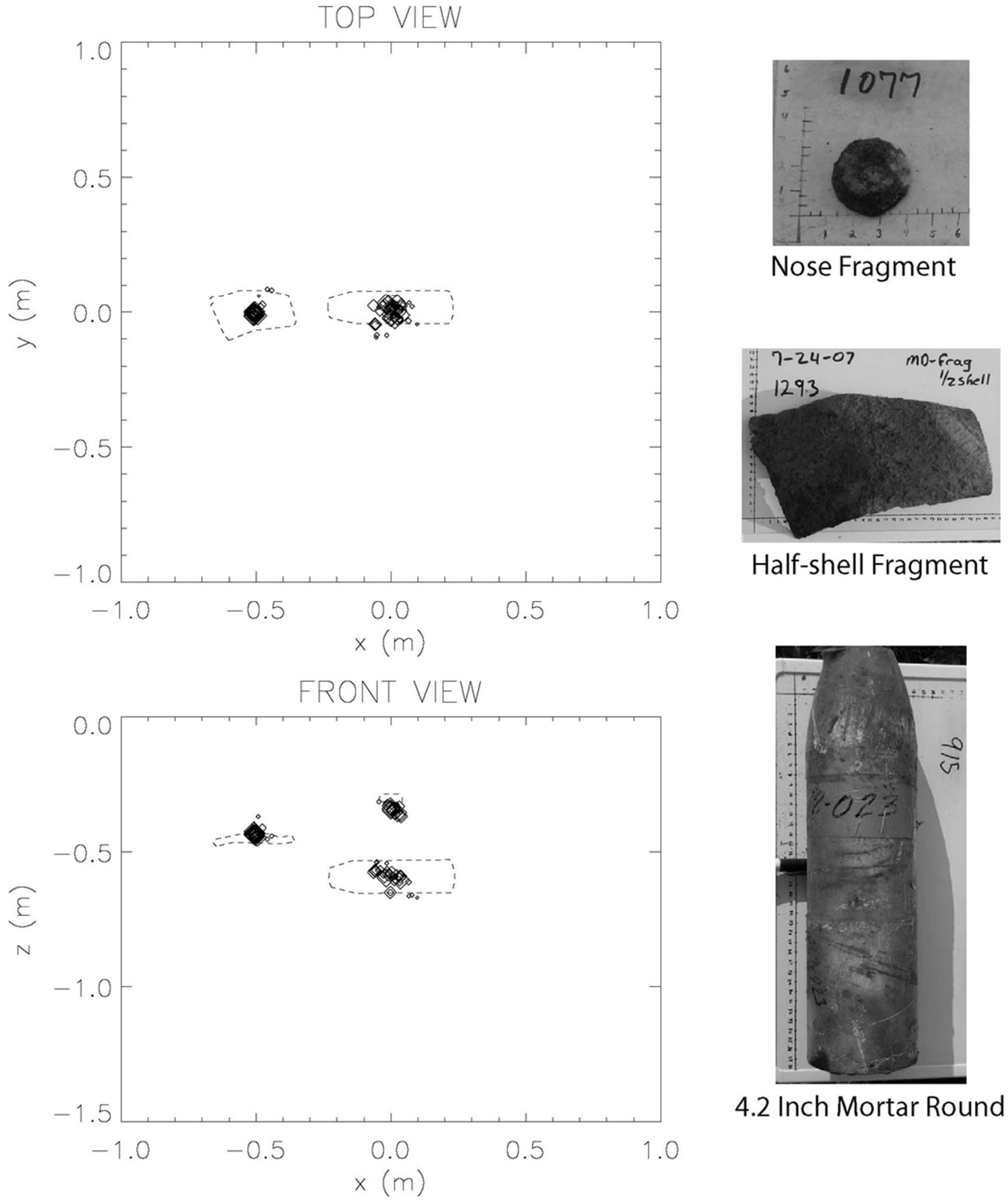
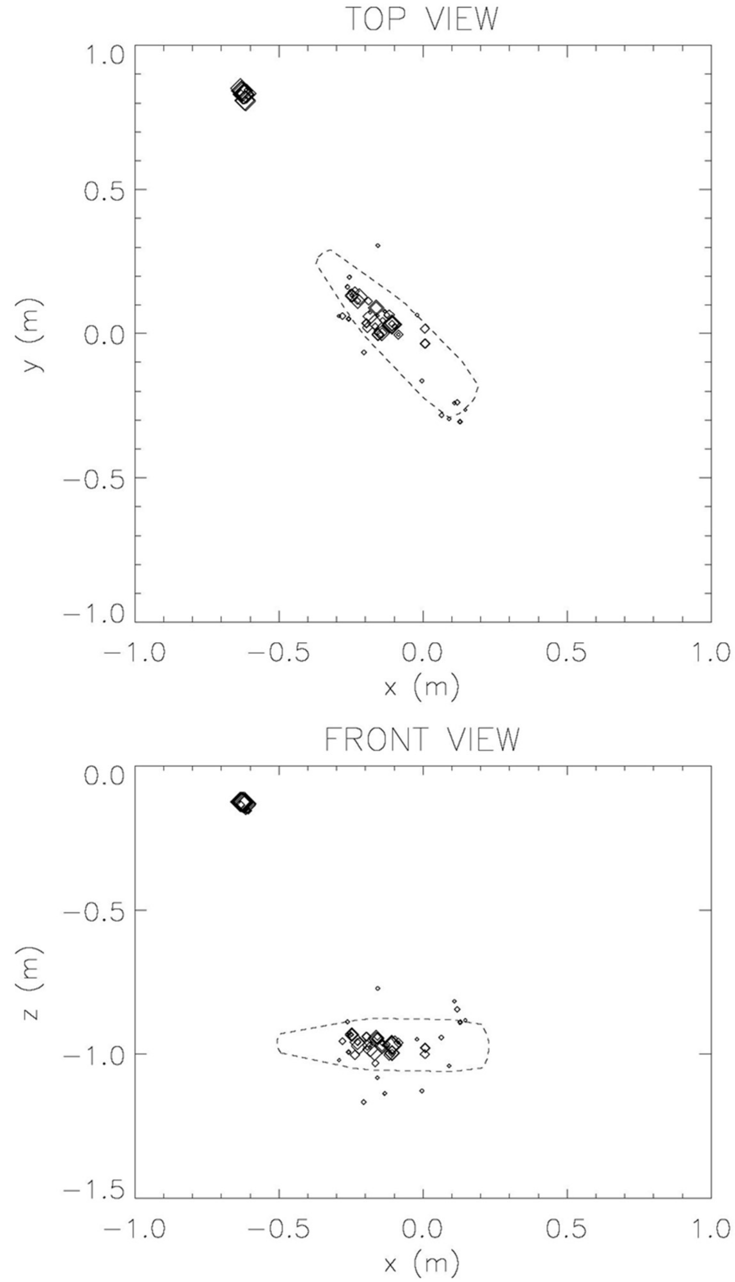


Figure 4. Results on data from Blossom Point collected in 2007 with the NRL TEM array. This test was conducted in air above ground with the sensor raised on non-conducting supports. The lowest object is a horizontal 4.2 inch mortar round, the next is a flattened “half-round” fragment, and the shallowest is a nose fragment. Black diamonds represent weights assigned to hypothetical sources by the algorithm, and the size of the diamond is proportional to the magnitude of the weight. Only the heaviest quartile of weights are shown. Dashed lines indicate true target locations. These results were obtained with the number of sources *p* set to 100, the size of the Harmony memory *k* set to 20, and the maximum number of iterations *nmax* set to 100. Processing time was under 20 seconds on a Dell PWS690 workstation.



This 155mm artillery round is similar to the one detected below ground at Aberdeen Proving Ground MD.

Figure 5. Results on data from Aberdeen Proving Ground (APG) collected in 2008 with the NRL TEM array. This is target A1 in the calibration grid at APG, which is a horizontal 155mm projectile centered approximately 0.7 meters below ground surface (z axis in these figures shows distance from the TEM sensor, not ground surface). As before, black diamonds represent weights assigned by the algorithm – heaviest quartile only - and the dashed line indicates the known target. The diamonds in the upper left corner are caused by an unknown shallow clutter item, apparently resting on ground surface. This clutter object complicates analysis of these data using standard single-target inversion methods to the point where the buried ordnance cannot be identified. These results were obtained in under 20 seconds on a Dell PWS690 workstation with parameters *p* =100*, k*=20, and *nmax*=100*,* as in figure 1

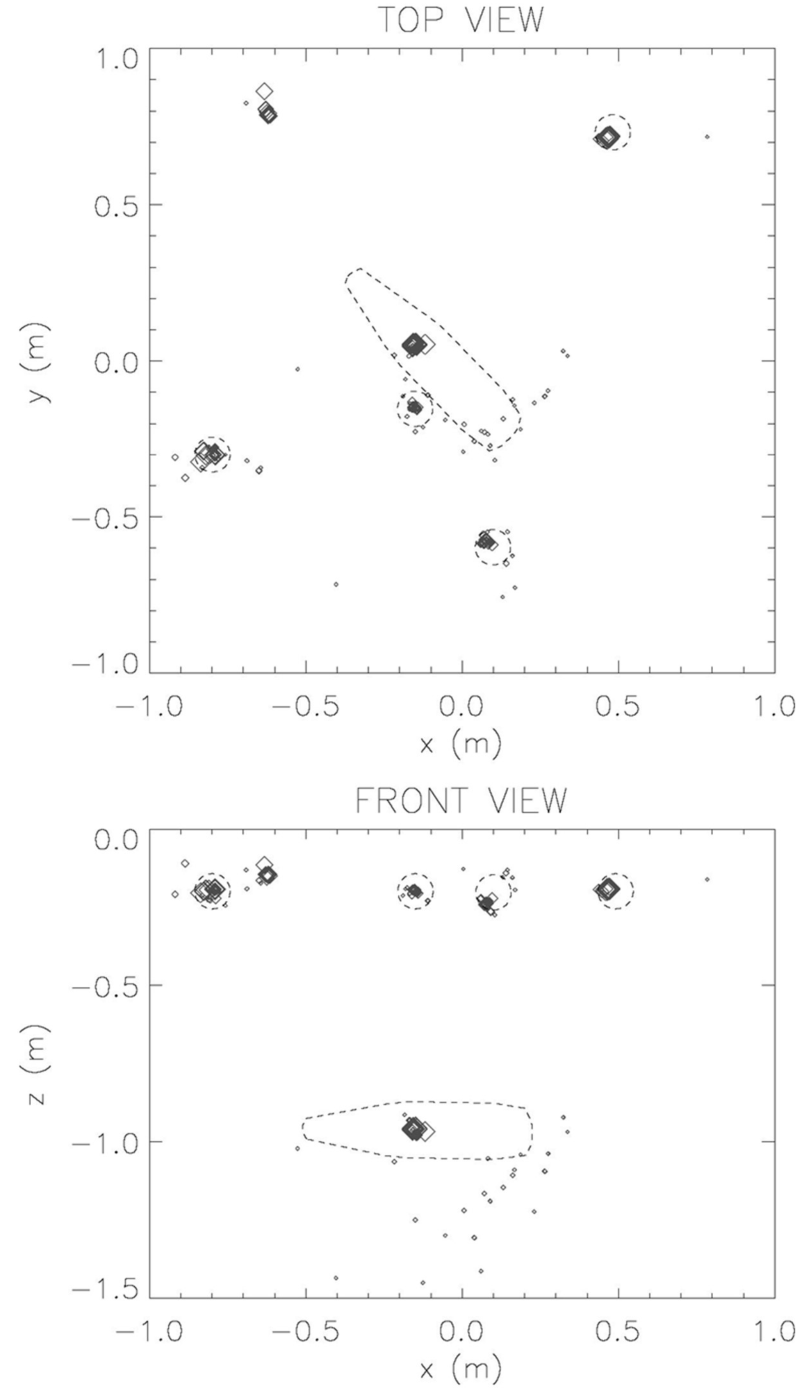


Figure 6. Four shallow synthetic sources were added into the APG data of figure 2. These four sources are spheres, all resting 20cm below the sensor, intended to represent near-surface clutter typically seen at active cleanup sites. The algorithm correctly located all six targets in this example, including the deep 155mm artillery round and the unknown shallow clutter in the upper left. Algorithm parameters were adjusted upwards somewhat in this example to allow for the more complicated environment: these were *p* =200*, k*=20, and *nmax*=400. Processing time was under 40 seconds. Again, only the heaviest quartile of weights are shown (black diamonds), and dashed lines represent known sources.

CONCLUSIONS AND IMPLICATIONS FOR FUTURE RESEARCH

An algorithm has been presented and demonstrated to locate and characterize multiple signal sources based on observation of their combined signals. The method relies on new solvers designed to provide sparse non-negative solutions for underdetermined linear systems of equations, permitting many hypothetical sources to be evaluated jointly with each iteration. Most of these sources receive zero weight, but those with positive weight drive successive refinement of the collection. A crucial consideration is the grouping of non-negative weights into local neighborhoods to estimate true sources which may be located inside. This grouping is done by partitioning the volume into tetrahedrons, each vertex of which is a non-zero source that received weight in the preceding iteration. Approximation of the presumed true source inside is accomplished by locating the tetrahedron centroid and then assigning source parameters to best match the summed signal from the four vertices. The overall number of sources in the system is indicated by the number of clusters in the output, which may be detected by clustering algorithms capable of counting clusters, such as Quality Threshold (QT) clustering.

Application to the problem of electromagnetic induction is demonstrated, using signals from the NRL TEM array. Results show successful location of multiple sources for both in-air test stand measurements and field measurements.

Future research is needed in speed improvements, automatic background removal, and the application to data acquired while moving.

LITERATURE CITED

David L. Donoho, Neighborly Polytopes and Sparse Solution of Underdetermined Linear Equations. http://sparselab.stanford.edu

David L. Donoho and Jared Tanner. Neighborliness of randomly-projected simplices in high dimensions. Technical report, Department of Statistics, Stanford University, 2005.

Barber, C.B., Dobkin, D.P., and Huhdanpaa, H.T., “The Quickhull algorithm for convex hulls”, ACM Trans. On Mathematical Software, 22(4):469-483, Dec 1996, http://www.qhull.org

Kang Seok Lee, and Zong Woo Geem. A new meta-heuristic algorithm for continuous engineering optimization: harmony search theory and practice. Computer Methods in Applied Mechanics and Engineering. Volume 194, issues 36-38, 23 September 2005. Pages 3902-3933.