FINAL REPORT

UXO Classification Demonstrations at Live Sites Using UX-Analyze

ESTCP Project MR-201312

APRIL 2020

Dean Keiswetter
Acorn SI

ESTCP

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
30/04/2020	ESTCP Final Report	4/25/2013 - 4/24/2018

4. TITLE AND SUBTITLE UXO Classification Demonstrations at Live Sites Using UX-Analyze	**5a. CONTRACT NUMBER** 13-C-0030
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S) Acorn SI: Dean Keiswetter Leidos: Thomas Bell	**5d. PROJECT NUMBER** MR-201312
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Acorn SI 205 Roebling Lane Carey, NC 27513	**8. PERFORMING ORGANIZATION REPORT NUMBER** MR-201312
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Environmental Security Technology Certification Program 4800 Mark Center Drive, Suite 16F16 Alexandria, VA 22350-3605	**10. SPONSOR/MONITOR'S ACRONYM(S)** ESTCP
	11. SPONSOR/MONITOR'S REPORT NUMBER(S) MR-201312

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

13. SUPPLEMENTARY NOTES

14. ABSTRACT
This report documents one demonstration in a series of Environmental Security Technology Certification Program (ESTCP) demonstrations of geophysical advanced classification technologies for munitions response (MR). This demonstration was designed to evaluate geophysical advanced classification methodology at a site with munitions types potentially ranging from a 20mm projectile to a 100 lb bomb. The Time-domain Electromagnetic Multi-sensor Towed Array Detection System MP 2x2 Cart (TEMTADS MP 2x2) system was demonstrated in both dynamic and cued modes at a site located on Andersen Air Force Base (AAFB), Guam.

15. SUBJECT TERMS
UXO, UXO Classification, Live Sites, UX-Analyze

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	41	**19b. TELEPHONE NUMBER** *(Include area code)* dkeiswetter@acornsi.com

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

FINAL REPORT
Project: MR-201312

TABLE OF CONTENTS

Page

Page

LIST OF FIGURES

Page

LIST OF TABLES

ACRONYMS AND ABBREVIATIONS

AFB	Air Force Base
DGM	Digital Geophysical Mapping
DPR MILCON	Defense Policy Review Initiative Military Construction
DSB	Defense Science Board
EM	electromagnetic
EMI	Electromagnetic Induction
EOD	Explosive Ordnance Disposal
ESTCP	Environmental Security Technology Certification Program
ft	feet
GTE	Ground truth evaluation
GPS	Global Positioning System
IDA	Institute for Defense Analyses
ISO	Industry Standard Object
IVS	Instrument Verification Strip
Large ISO40	Large Industry Standard Object 40 = 4″ x 12″ schedule 40 pipe nipple
MEC	munitions and explosives of concern
Medium ISO40	Medium Industry Standard Object 40 = 2″ x 8″ schedule 40 pipe nipple
MR	munitions response
MRS	Munitions Response Site
NRP	North Ramp Parking
Pclass	Probability of Correct Classification of TOI
QC	quality control
RMS	root mean square
SAC	Strategic Air Command
SERDP	Strategic Environmental Research and Development Program
Small ISO80	Small Industry Standard Object 80 = 1″ x 4″ schedule 80 thick wall pipe nipple
SNR	signal to noise ratio
TEMTADS	Time-domain Electromagnetic Multi-sensor Towed Array Detection System
TOI	Target of Interest
U.S.	United States

UTM	Universal Transverse Mercator
UXO	Unexploded Ordnance
VSP	Visual Sample Plan
WESTON®	Weston Solutions, Inc.
WWII	World War II

1.0 INTRODUCTION

This report documents one demonstration in a series of Environmental Security Technology Certification Program (ESTCP) demonstrations of geophysical advanced classification technologies for munitions response (MR). This demonstration was designed to evaluate geophysical advanced classification methodology at a site with munitions types potentially ranging from a 20mm projectile to a 100 lb bomb. The Time-domain Electromagnetic Multi-sensor Towed Array Detection System MP 2x2 Cart (TEMTADS MP 2x2) system was demonstrated in both dynamic and cued modes at a site located on Andersen Air Force Base (AAFB), Guam.

The TEMTADS MP 2x2 demonstration data collection was performed in the North Ramp Parking (NRP) area of AAFB (Figure 1-1) by Weston Solutions, Inc. in January and February 2014. This document summarizes the advanced classification analysis of TEMTADS MP 2x2 data performed by NAEVA Geophysics, Inc. (NAEVA) on data collected as part of the ESTCP Munitions Response Live Site Demonstration at the North Ramp Parking (NRP) area of AAFB, Guam.

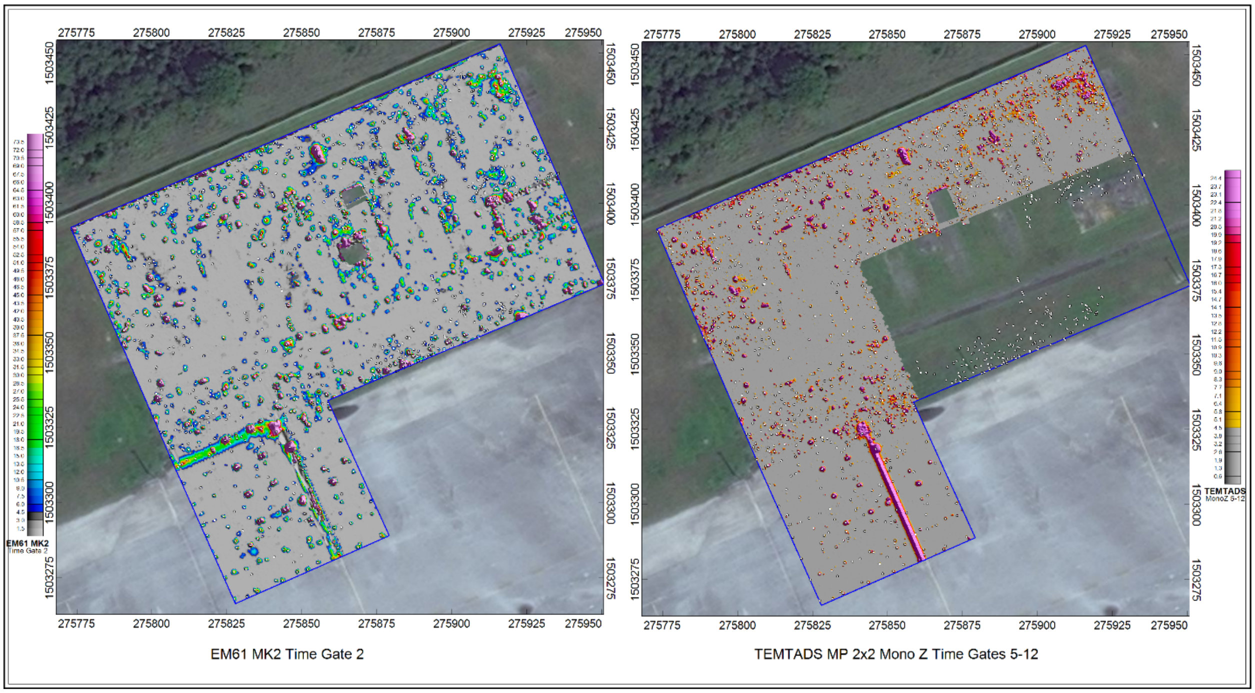

Figure 1-1. Dynamic detection survey data at the NRP area of AAFB, Guam.

The objective of the demonstration was to validate geophysical classification technology at AAFB. NAEVA will perform advanced classification of the cued geophysical data using the UX-Analyze module of Geosoft's Oasis montaj software.

2.0 TECHNOLOGY

2.1 TECHNOLOGY DESCRIPTION

This demonstration consisted of the classification analysis of TEMTADS MP 2x2 data with the UX-Analyze module of Geosoft's Oasis montaj software. A testing version of the software, v8.4.0.1222 Testing, was made available for this demonstration to make use of the latest developments in streamlined cued classification workflow.

The UX-Analyze Advanced module is a commercially available solution for the processing and analysis of advanced sensor data collected with the MetalMapper and TEMTADS systems. The module includes import functions, QC tools, data modeling routines, IVS tools, library management functions and classification tools. Additional data handling and performance assessment functions are also available.

2.2 ADVANTAGES AND LIMITATIONS OF THE TECHNOLOGY

The major advantage of advanced EMI sensors and UX-Analyze software is that combined, they provide the ability to classify anomalies as being either TOI or non-TOI. Conventional DGM sensors (e.g., EM61-MK2) have very limited ability to discriminate between TOI and non-TOI. The MP System was originally designed to offer similar cued-mode production rates to those seen for larger, vehicular-towed advanced EMI sensors while providing the ability to operate in difficult terrain and treed areas that the larger systems cannot access. With the upgraded TEMTADS/3D sensors, similar performance was achieved with similar classification-grade data quality. The MP array is 80 cm on a side and mounted on a man-portable cart.

There is a limiting anomaly density above which the response of individual targets cannot be successfully differentiated. Small sensors were chosen for this array, which helps mitigate this problem but cannot eliminate it completely. Recent developments, including solvers designed for classification in multiple-object scenarios such as SAIC's multi-target solver, are being evaluated and their performance characteristics in cluttered environments determined.

3.0 PERFORMANCE OBJECTIVES

The performance objectives are summarized in Table 3-1. These objectives apply to all advanced classification work performed on the NRP TEMTADS MP 2x2 dataset.

Table 3-1. Performance Objectives

Performance Objective	Metric	Data Required	Success Criteria
Repeatability of IVS Measurements	Library match and target location	The first day's IVS measurement and seed location	Qualitative evaluation for repeatability and noise variations throughout project
Correctly Identify All Seeds and Native TOI	All Seeds and All Native TOI	Percent classified as TOI	100% classified as TOI
Correctly Identify Size Group	All targets above the dig threshold	Percent of targets above the dig threshold grouped correctly	85% correctly grouped by size
Correct estimation of target location	All excavated anomalies	Measured and predicted location	ΔX, $\Delta Y < 15$cm (1σ)
Maximize correct classification of Non-TOI	All Non-TOI	Number of false alarms eliminated	Reduction of clutter digs by >50% while meeting all other demonstration objectives
Minimize number of anomalies that cannot be analyzed	All cued anomalies	Number of anomalies that must be classified as "Unable to Analyze"	Reliable target parameters can be estimated for > 95% of anomalies on each sensor's detection list.

3.1 OBJECTIVE: REPEATABILITY OF IVS MEASUREMENTS

The reliability of the survey data depends on the proper functioning of the survey equipment. This objective concerns the twice-daily confirmation of sensor system performance.

3.1.1 Metric

The metrics for this objective are qualitative due to the analysis being performed well after data collection has been completed. The inverted polarization library match and target location were evaluated against a standard library and a self-match to determine repeatability.

3.1.2 Data Requirements

The IVS data will be used to judge this objective. A standard library with a wide range of munitions items as well as the modeled IVS data were used.

3.1.3 Success Criteria

The objective will be considered met by generating a summary table of IVS statistics to be used to guide preliminary threshold selections during the classification process.

3.2 OBJECTIVE: CORRECTLY IDENTIFY ALL SEEDS AND NATIVE TOI

This metric applies to QC seeds, population seeds, and native TOI. Seed items are used to provide objective and quantitative measurement of the classification process and are used to supplement advanced classification objectives.

The seeds for this demonstration are small ISO80, 20mm, 25mm, 37mm, and 47mm. The objective for the advanced classification process for this demonstration is to correctly classify 100% of all TOI.

3.2.1 Metric

The metrics for this objective are the percentage of TOI correctly identified on the TOI lists.

3.2.2 Data Requirements

Ranked anomaly lists, separated into TOI and non-TOI lists, are used to judge the success of this objective.

3.2.3 Success Criteria

The objective will be considered a success if 100% of the QC seeds, population seeds, and native TOI are placed on the TOI list.

3.3 OBJECTIVE: CORRECTLY IDENTIFY SIZE GROUP

The demonstrators will attempt to correctly classify each TOI as small, medium, or large. The groups are designated as follows:

- Small group (<50mm diameter).
- Medium group (50mm to <100mm diameter).
- Large group (≥100mm diameter).

3.3.1 Metric

The metrics for this objective are the percentage of anomalies above the stop dig point correctly classified as small, medium, or large.

3.3.2 Data Requirements

Anomalies grouped as small, medium, or large are used to judge the success of this objective. Grouping the anomalies accurately may depend on the usability of the $\beta 2$ and $\beta 3$ polarizability curves. If the quality of the curves cannot support the demonstrator's methodology, the demonstrator will identify the anomaly as "can't make a size determination" and provide an explanation.

3.3.3 Success Criteria

The group assignment task will be considered a success if 85% or more of the anomalies above the stop dig point are grouped correctly.

3.4 OBJECTIVE: CORRECT ESTIMATION OF TARGET LOCATION

This objective involves the accuracy of the target fit location in the first phase of the analysis (data inversion.) Accurate target location estimation is vital for the intrusive operations and the more reliable the predicted locations are the greater the likelihood the correct object will be recovered.

3.4.1 Metric

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

3.4.2 Data Requirements

The estimated parameters will be compared to those measured during the intrusive investigation. For targets with multiple sources, only the highest ranked fit result will be evaluated.

3.4.3 Success Criteria

The objective will be considered to be met if the estimated X, Y locations are within 15 cm.

3.5 OBJECTIVE: MAXIMIZE CORRECT CLASSIFICATION OF NON-TOI

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

Because the number of clutter items that may resemble 20mm projectiles is unknown the success metric for this objective (50%) is lower than most previous demonstrations, which typically use a metric of 65%.

3.5.1 Metric

The metric for this objective is the number of cued anomalies that can be correctly classified as Non-TOI.

3.5.2 Data Requirements

Each demonstrator prioritized Non-TOI on the cued anomaly list. IDA personnel will use their scoring algorithms to assess the results.

3.5.3 Success Criteria

The objective will be considered to be met if more than 50% of the Non-TOI items can be correctly labeled as Non-TOI while meeting the objectives or success criteria for TOI stated in Table 3-1.

3.6 OBJECTIVE: MINIMIZE NUMBER OF CANNOT ANALYZE ANOMALIES

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

3.6.1 Metric

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

3.6.2 Data Requirements

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

3.6.3 Success Criteria

The objective will be considered to be met if reliable parameters can be estimated for > 95% of the anomalies on each sensor anomaly list.

4.0 SITE DESCRIPTION

The NRP demonstration area is scheduled to undergo construction and due to known or suspected MEC DPRI MICLON program requires a removal action in advance of construction. The Weston portion of the demonstration was integrated with the previously scheduled MEC removal action. Results of the demonstration provide insight and consideration for future use of geophysical classification ahead of DPRI MILCON programs and projects. Also, besides the DRPI MILCON efforts, there are a number of Military Munitions Response Program sites on Guam that will benefit from the results of this demonstration project.

4.1 MUNITIONS CONTAMINATION

The expected munitions at Andersen AFB NRP area are listed in Table 4-1, which are based on the information in the Explosive Safety Submission [1] and the current conceptual site model. The known items in the table identify munitions types recovered by explosive ordnance disposal (EOD) Teams between 1991 and 2011, 20mm are suspected to be at the site however have not been recovered.

Table 4-1. Known and Suspected Munitions Types

Andersen AFB, NRP Area
MK II Hand Grenades
20mm projectile
60mm mortar
81mm mortar
105mm projectiles
155mm projectile
5-inch projectile
6-inch projectile
100 lb bomb

In addition to these munitions the Weston seeding plan states that the QC and population seeds used for this demonstration were small ISO80, 20mm, 25mm, 37mm, and 47mm.

5.0 TEST DESIGN

The objective of this program is to demonstrate commercially available technology for the use of geophysical classification in the munitions response process. The key component of this methodology is the analysis of the selected anomalies using physics-based models to extract target parameters such as size, shape, and materials properties; and the use of those parameters to construct a ranked anomaly list.

WESTON provided cued data for analysis and detection survey data for reference. The cued data set was processed using existing UX-Analyze routines to extract target parameters. These parameters were passed to the classification routines that, after training on a limited amount of site-specific ground truth, were used to produce ranked anomaly lists. These inputs were scored by IDA with emphasis on the number of items that were correctly labeled non-hazardous while correctly labeling all TOI.

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

5.1 CALIBRATION ACTIVITIES

The IVS survey and test pit data were qualitatively evaluated as calibration activities to base preliminary classify and rank threshold values. Factors such as the repeatability of the measurements over time, fit coherence and confidence metric at the IVS were evaluated. For the Test Pit any measurements that represented unique signatures were incorporated into the initial signature library.

5.2 VALIDATION

At the conclusion of data collection and classification activities, all anomalies on the master dig lists for the NRP area were excavated by the removal action contractor. Each item encountered was identified, its depth measured, its location determined using cm-level GPS, and the item removed if possible.

6.0 DATA ANALYSIS AND PRODUCTS

All data analysis was performed using the UX-Analyze Advanced module of Oasis Montaj. The IVS and test pit data were processed to verify consistent instrument response and populate the site specific classification library as necessary. NAEVA followed the general outline of the cued analysis process described in section 6.0 of the Weston Demonstration Plan, an overview of the workflow is shown in Figure 6-1.

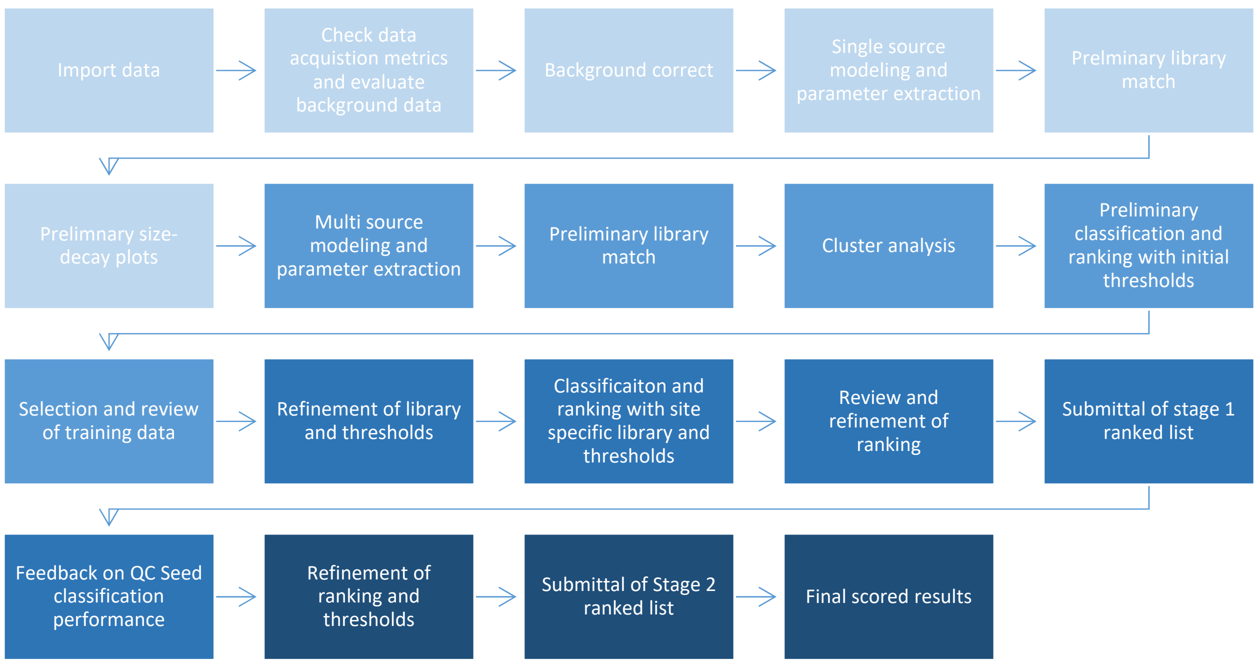

Figure 6-1. Overveiw of Data Analysis Workflow.

Main steps included data verification, parameter extraction, library validation and training, and classification and ranking.

Data verification was performed to evaluate background data, GPS, IMU, and transmitter current. All cued data were background corrected in preparation for parameter estimation. Single and multi-solver algorithms were used to invert the data and extract parameters. Preliminary library classification and cluster analysis were performed and training data were requested as needed for library validation to build a site specific classification library.

Data went through a classification process and ranked lists were submitted to the ESTCP Program Office for scoring. Three independent analysts submitted ranked lists for scoring. Data provided for analysis included processed dynamic TEMTADS data; classification incorporated the use of the dynamic data, however only used target parameters extracted from the cued data.

This summary report was prepared to document the site-specific library and classification decision process and includes a performance assessment. ROC curves constructed by IDA are included along with summary statistics to assess the Performance Objectives.

6.1 DATA VERIFICATION

Although data collection was completed prior to the start of data analysis by NAEVA, the data verification process in UX-Analyze was performed to evaluate the data prior to parameter estimation. All survey data were evaluated for GPS quality, IMU functionality and transmitter current statistics to ensure usability. The background measurement database was evaluated using the Calculate Background Database QC Stats tool and Decay Display – QC. Any survey and background data with suspect data quality or positional errors were marked and not used for classification.

After QC checks were performed the survey data were background corrected in preparation for parameter estimation. The final stage of data verification was performed with the "Field Modelling Check" routine and included single source modeling of the data and a preliminary library match to a comprehensive library containing examples of expected TOI at the site as well as a range of other MEC items. Decay plots, data maps, polarization plots, and size and decay maps were created and evaluated as part of this routine.

6.2 PARAMETER ESTIMATES

All anomaly survey data were modeled with single and multi-solver inversion methods to extract parameters. The single solver modeled the data for a single target as the anomaly source and saved the extracted features to a line in the Targets database. The multi-solver in UX-Analyze used the sparse array solver method which generated multiple realizations of solutions for each data collection event. The number of realizations can be customized by the user by changing the number of restarts to the process, each restart will generate a set of eight realizations that use different clustering parameters to identify sources during modeling. For the NRP both analysts used 5 restarts resulting in 40 realizations for each cued reading. All realizations were saved for use during classification and the extracted features for the current (or active) solution were saved to the Targets database.

6.3 LIBRARY VALIDATION AND TRAINING

Following a preliminary classification and ranking of the data, along with cluster analysis performed using the "Validate Library" routine in the UX-Analyze module, preliminary decision plots were created which include scatter plots, the multi-solver fit result, the library match for the top multi-solver fit result, and the cluster polarization plot. The interactive size/decay cluster tool was used to further interrogate cluster groupings.

A number of anomalies were selected from the ranked list to better characterize on-site TOI and refine classification thresholds, and from representative samples from identified unique clusters to assist in refinement of the generic library to a site-specific library. Unnecessary entries from the generic library were excluded and any additional TOI or unique clutter identified from the training data request were incorporated.

6.4 CLASSIFY AND RANK

Data for each anomaly were classified using the "Classify and Rank" routine in the UX-Analyze module. This routine combines library matching with rule-based decision making.

The library matching classification process was performed following any updates to the library and the rule-based ranking was performed several times throughout the entire decision-making process followed by an analyst review. The Classification tool applies user-defined thresholds and performs library matching for all modeled data using several different parameter metric weights. During classification all multi-solver realizations are examined and the current solutions in the Targets database are updated with the realization with the best match to a library entry. All ranked fit results were evaluated by the processors and targets were re-ranked and/or re-classified at processor discretion. This included the substitution of alternate fit results, the reclassification of targets with a good visual match not fully captured by the combined confidence metric, or the reclassification of targets based on their location in Size-Decay parameter space. Data review by the analyst included evaluation of the components of the data displayed in the example decision plot shown in Figure 6-2, as well as dynamic data collected over the anomaly, the best 3-criteria match to clutter identified from requested ground truth, and multisource polarization plots.

.

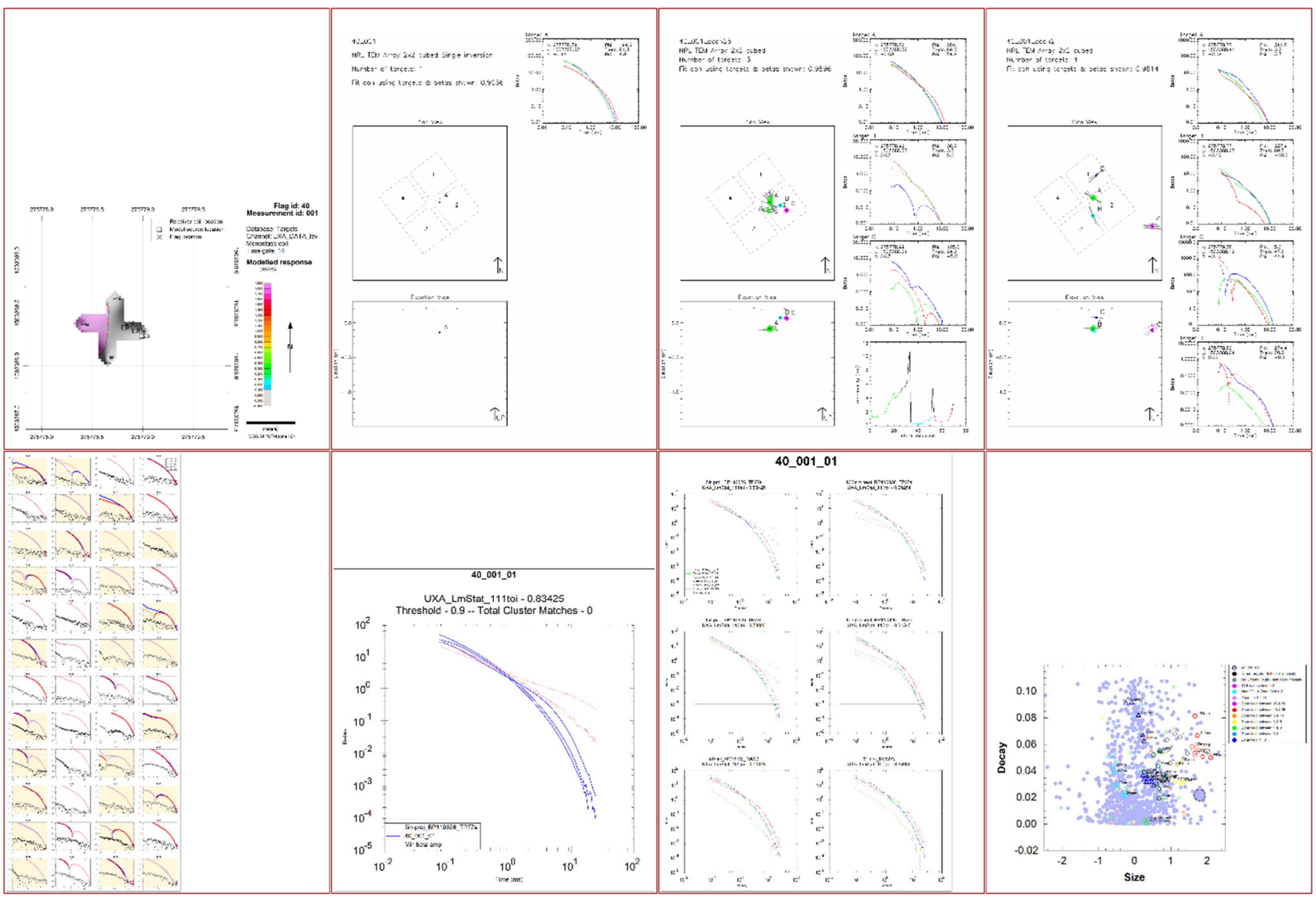

Figure 6-2. Decision Plot for a Cued Target Displaying (Clockwise from Upper Left) Cued TEMTADS Survey Data with Target and Fit Locations, Single Source Inversion Result, Multisource Inversion Result, Multi_110 Source Inversion Result, Size/Decay Plot, Library Match Polarizability Plot, Cluster Polarizability Plot, Transient Decay Traces

6.5 DECISION MEMO

Thresholds were assigned for several parameters as part of the classification process. Threshold values were applied using the "Set thresholds and prioritize" tool contained in the "Classify and Rank" routine. The TOI threshold was determined based on a substantial drop-off in the visual similarity between the extracted polarizabilities and library TOI, training data ground truth, and intrusive results from the top ranked targets on the dig list.

Using the parameters outlined above, NAEVA produced ranked lists classifying each target from the processed data set. The first items on the initial ranked anomaly list are those targets for which ground truth had been requested. Following this, anomalies for which reliable parameters could not be extracted and therefore must be dug are listed. Next is the item deemed most likely a TOI followed by all items that are possibly TOI. Finally, all those items that the demonstrator was confident are Non-TOI are ranked by their confidence (least likely Non-TOI at the top, most likely Non-TOI at the bottom.) In addition, NAEVA provided an expected diameter range for the source (less than 50mm, between 50mm and 100mm, and greater than 100mm) for targets above the dig threshold.

6.5.1 Ground truth evaluation 1 (GTE1)

Training data were requested for three purposes; to identify sources that were not present on site and could be removed from the site-specific library used for final classification, to investigate preliminary buffers, and to investigate source items from clusters identified during the library validation stage. During the library validation stage, default values were used for all thresholds, with the exception of the cluster library threshold, the library subset threshold, and minimum size. These were adjusted to correspond to the preliminary Buffer and TOI and Buffer and Non-TOI weak thresholds and to represent the minimum size for the smallest expected target of interest recorded in the classification library, respectively. Ground truth data was used to evaluate and update the Buffer and TOI and Buffer and non-TOI thresholds. Evaluation of IVS data collected over the duration of the demonstration lead to a Buffer and Non-TOI weak threshold of 0.8 for all targets with a size ($\beta 1$) library match below 0.94 and signal amplitude below 15. The default of 1 was used as the cluster library threshold to identify similar items during classification. A small subset of targets was re-categorized at processor discretion based on review of the data. Targets were designated as "Cannot Analyze" if any of the following are true for all of the best-match inversion results: the inversion process did not finish, the extracted fit depth was below the designated threshold, two or more of the polarizabilities were negative, the fit coherence was below threshold and/or the difference between the array and the source was greater than the set threshold.

Table 6-1. GTE1 Selected Thresholds

	Preliminary	Final
Gates (Low/High/Decay)	14/120/71	14/120/71
Chi2 Threshold	50	50
Cluster Self Match Threshold	0.9	0.9
Min Beta Amplitude	0.001	0.001
Cluster Library Threshold	0.925	1
Library Subset Threshold	0.75	N/A
Cluster Min # Matches	3	3
Signal Amplitude (Min/Max)	2/20	2/15
Fit Depth	2	2
Decay (Min/Max)	0/0.138	0/0.138
Size (Min/Max)	-0.65/3.144	-0.65/3.144
D Array-Source	0.4	0.4
D Array-Flag	0.75	0.75
D Flag-Source	0.6	0.6
Fit Coherence	0.8	0.8
Buffer and TOI (Category 1)	0.925	0.95
Buffer and Non-TOI (Category 3)	0.825	0.9
Buffer and Non-TOI weak (Category 2)	0.75	0.8
High Confidence Clutter	0.925	0.95

Table 6-2. GTE1 Ranking Expression Variables

	Description
t	Training
x	Cannot Analyze
c	Likely Clutter
uc	Cannot Classify
u	Likely TOI

6.5.2 Ground truth evaluation 2 (GTE2)

Ground truth was requested for targets identified during the library validation stage as having a unique signature based on self-match clustering, as well as for targets with a high number of cluster matches to TOI in the library but only moderate library match statistics. A cluster library threshold was matched to the default buffer and TOI metric, but otherwise default values were used for library validation. Based on evaluation of the training data and library validation, the initial library was refined to most likely only contain site-specific items, which included adding data for clutter items to the library. Targets were designated as “Cannot Analyze” based on the same criteria listed in section 6.5.1. All targets were evaluated by the analyst and their priority adjusted on the ranked list based on visual inspection of the dynamic data, relation to nearby targets, signal amplitude, single and multi-solver polarizability curves, and other target parameters.

Table 6-3. GTE2 Selected Thresholds

	Preliminary	Final
Gates (Low/High/Decay)	14/120/71	14/120/71
Chi2 Threshold	50	50
Cluster Self Match Threshold	0.9	0.9
Min Beta Amplitude	0.001	0.001
Cluster Library Threshold	0.925	1
Cluster Min # Matches	3	3
Signal Amplitude (Min/Max)	2/20	2/20
Fit Depth	2	2
Decay (Min/Max)	0/0.138	0/0.138
Size (Min/Max)	-0.33/3.144	-0.33/2.812
D Array-Source	0.4	0.4
D Array-Flag	0.75	0.75
D Flag-Source	0.6	0.6
Fit Coherence	0.8	0.8
Buffer and TOI (Category 1)	0.925	0.95
Buffer and non-TOI (Category 3)	0.825	0.85
Buffer and non-TOI weak (Category 2)	0.75	0.75
High Confidence Clutter	0.925	0.925

Table 6-4. GTE2 Ranking Expression Variables

	Description
t	Training
x	Cannot Analyze
c	Likely Clutter
uc	Cannot Classify
u	Likely TOI

6.5.3 Ground truth evaluation 3 (GTE3)

Training data were selected to evaluate unique clusters identified during the validate library stage with self-match, identify unique non-clustering targets, and evaluate large clusters with moderate library match statistics. Default values were selected for the validate library stage, except for the cluster library threshold which was matched to the default buffer and TOI metric. Results from the ground truth aided in refining the library to remove items unlikely to be encounter on the site. A subset of ground truth was also added as a clutter line to the library. Initial ranking was performed using default values. Targets were assigned to the "Cannot Analyze" category using the same criteria as in sections 6.5.1 and 6.5.2. Visual inspection of targets was performed by the analyst to make a number of manual rank adjustments and to refine the initial default rank thresholds. Elements considered during analyst inspection included target signal amplitude, fit coherence, noise, single and multi-solver polarizability curves, dynamic data appearance and spatial relation to adjacent targets, alternative multi-solver solutions, and additional target parameters.

The thresholds for Buffer and TOI, Buffer and non-TOI, and Buffer and non-TOI weak targets were adjusted to 0.94, 0.895, and 0.84, respectively from initial values.

Table 6-5. GTE3 Selected Thresholds

	Preliminary	Final
Gates (Low/High/Decay)	14/120/71	14/120/71
Chi2 Threshold	50	50
Cluster Self Match Threshold	0.9	0.9
Min Beta Amplitude	0.001	0.001
Cluster Library Threshold	0.925	1
Cluster Min # Matches	3	3
Signal Amplitude (Min/Max)	2/20	2/20
Fit Depth	2	2
Decay (Min/Max)	0/0.138	0/0.138
Size (Min/Max)	-0.33/3.144	-0.33/3.144
D Array-Source	0.4	0.4
D Array-Flag	0.75	0.75
D Flag-Source	0.6	0.6
Fit Coherence	0.8	0.8
Buffer and TOI (Category 1)	0.925	0.94
Buffer and non-TOI (Category 3)	0.825	0.895
Buffer and non-TOI weak (Category 2)	0.75	0.84
High Confidence Clutter	0.925	0.925

Table 6-6. GTE3 Ranking Expression Variables

	Description
t	Training
x	Cannot Analyze
c	Likely Clutter
uc	Cannot Classify
u	Likely TOI

7.0 PERFORMANCE ASSESSMENT

The performance objectives for the demonstration are given in section 3.0. The results of these objectives are summarized in Table 7-1. Details of the data analysis in support of the assessment of performance objectives is provided in the following subsections.

Table 7-1. Performance Assessment

Performance Objective	Metric Success Criteria	GTE1	GTE2	GTE3
Repeatability of IVS Measurements	Qualitative evaluation for repeatability and noise variations throughout project	The first day's IVS measurement and seed location		
Correctly Identify All Seeds and Native TOI	Classify 100% of seeds and native TOI as TOI	100% TOI	100% TOI	100% TOI
Correctly Identify Size Group	Correctly group 85% of targets above the dig threshold by size	See Section 7.2.1	See Section 7.2.2	See Section 7.2.3
Correct estimation of target location	ΔX, ΔY < 15cm (1σ) for all excavated anomalies	18% Within 15cm	17% Within 15cm	17% Within 15cm
Maximize correct classification of Non-TOI	Reduction of clutter digs by >50% while meeting all other demonstration objectives	79% Non-TOI below stop dig line	95% Non-TOI below stop dig line	68% Non-TOI below stop dig line
Minimize number of anomalies that cannot be analyzed	Reliable target parameters can be estimated for > 95% of anomalies on each sensor's detection list.	2.7% Cannot Analyze	0.6% Cannot Analyze	5.4% Cannot Analyze

7.1 IVS

The IVS was processed following the data verification and parameter estimation processes described in sections 6.1 and 6.2. The modeled features from the initial survey were used for the library to evaluate the subsequent surveys.

7.1.1 Evaluation of Effect of Equipment Change

There was a change in the TEMTADS MP 2x2 sensor part way through the project. Overlays of background and raw IVS data from before and after the change show different responses from the instrument. An example of the monostatic decays for background readings before and after the change are shown in Figure 7-1; the 1Z_1Z and 1Z_1Y traces show significant differences in response and decay. After background corrections are applied and the data is modeled the extracted parameters are consistent throughout the project. It was determined that classification of the data would not require different consideration of data collected before and after the modifications to the equipment.

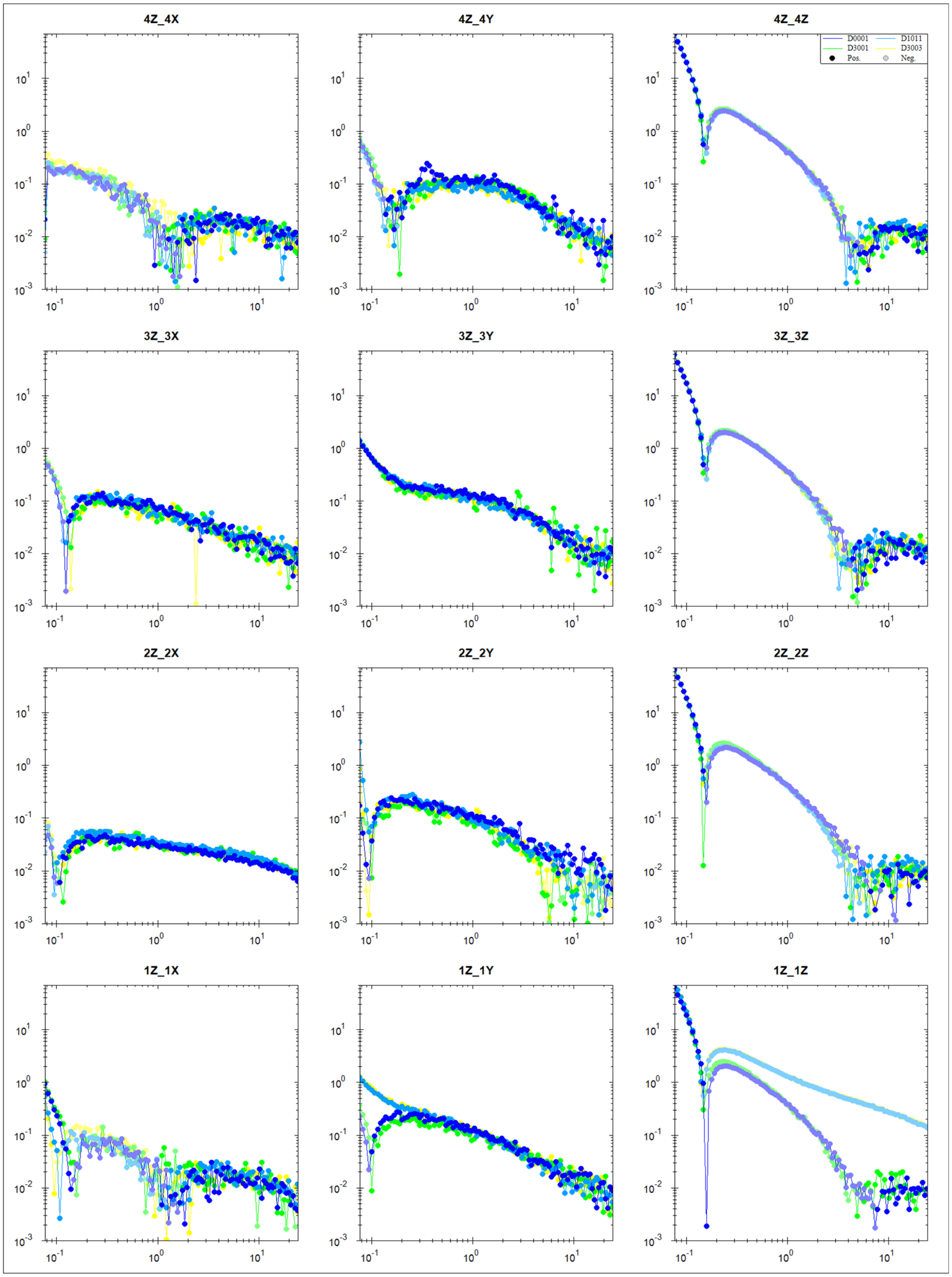

Figure 7-1. Example of Monostatic Background Readings Before and After Equipment Change.

7.1.2 Evaluation of Fit Coherence

Generally, the fit coherence was good, with most results over 0.95. Based on the size of the seed items and the depths to which they were buried it is expected that seed items and native TOI to the depths expected at the site should be able to be modeled with a reasonable fit coherence when good quality data is collected.

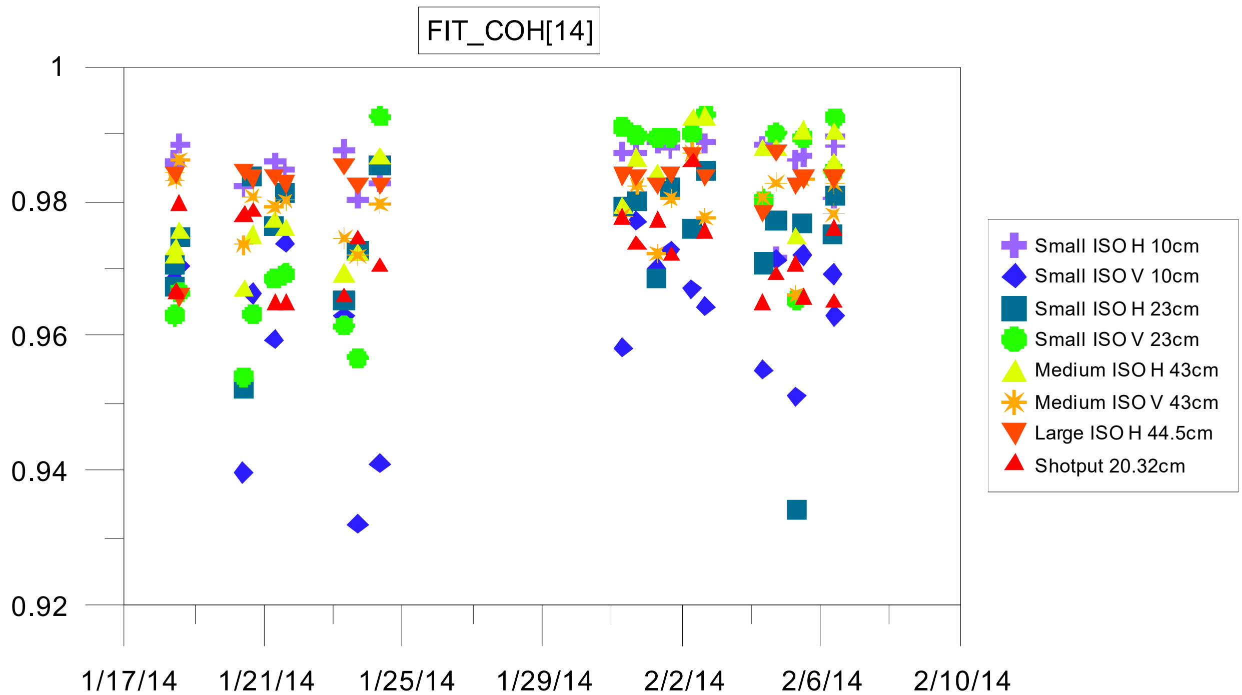

Figure 7-2. Fit coherence from Single Solver Inversion of IVS Measurements.

7.1.3 Evaluation of Confidence Metric

The confidence metric using a three criteria equal weight self-match to the first day's IVS measurements was performed and the best library match for each measurement was evaluated. For smaller and deeper test items there is a significant decrease in the confidence metric. The extracted betas for these items exhibit higher noise levels as well as overall variation in the amplitude, shape/decay and ratios between betas, all of which contribute to a lower confidence metric. This drop-off in confidence metric for lower signal amplitude targets was taken into account when establishing the Buffer and Non-TOI weak threshold for data classification.

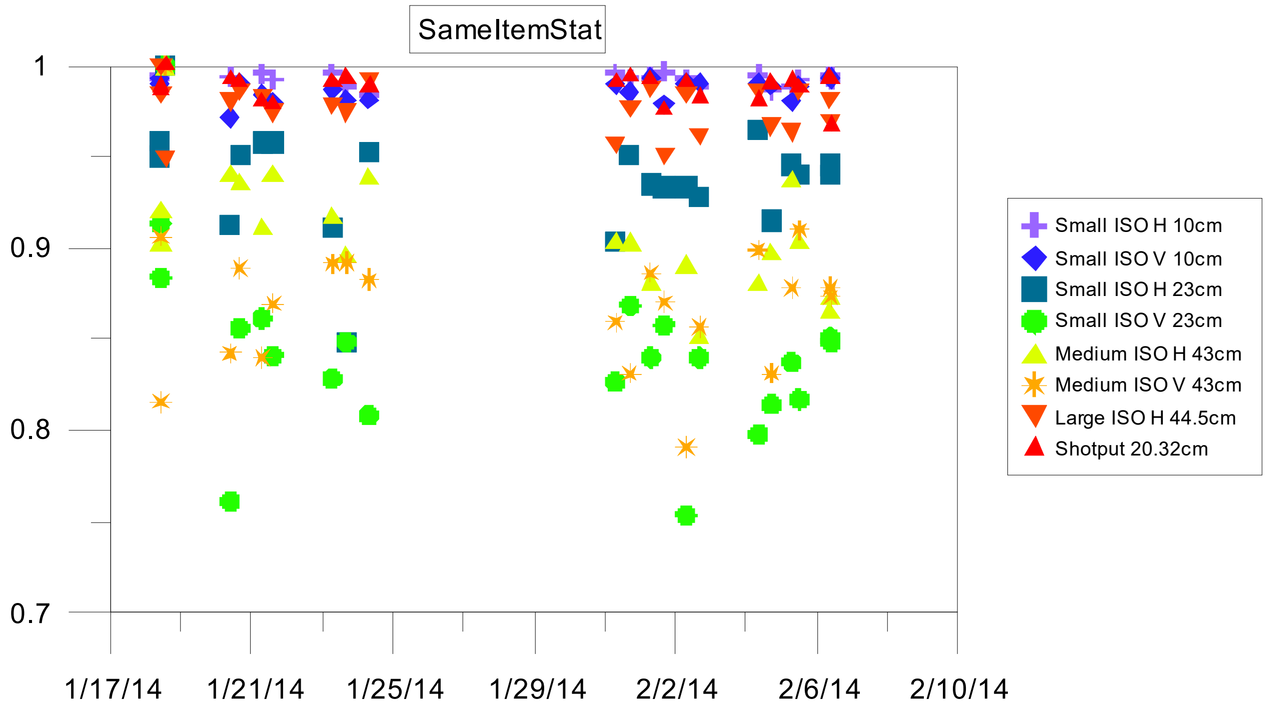

Figure 7-3. Three Criteria Confidence Metric for all IVS Measurements.

7.2 SCORED CLASSIFICATION RESULTS

NAEVA's ranked anomaly lists were scored against the emplaced and recovered targets and were analyzed to confirm that all QC seeds had been classified correctly. The results and ROC curves for the two independent classification analysts are shown in figures 7-4 and 7-5. The key regions to interpret the curves used in this program are:

- Origin to Point A: Targets for which ground truth was provided to aid in classification.
- Point A to Point B: Targets requiring intrusive investigation because they cannot be classified.
- Point C: Location of the lowest ranked TOI on the dig list.
- Point D: NAEVA's stop dig point, the threshold for the dividing point between targets classified as potential TOI and Non-TOI.
- Point D to Point E: Targets that were below the stop dig point representing the reduction in intrusive effort gained through classification. For successful classification, all targets of interest will fall to the left of Point D and the majority of non-TOI will be located to the right of point D.

The primary performance metric is the point at which the curve reaches 100% identification of TOI. The number of non-TOI correctly classified is a measure of the savings possible through classification.

7.2.1 GTE1

All seed items were correctly classified as high confidence TOI. No native items were present within the demonstration area. 246 Non-TOI items remained above the dig line, resulting in 931 Non-TOI that did not have to be dug. Classification results are summarized in figure 7-4.

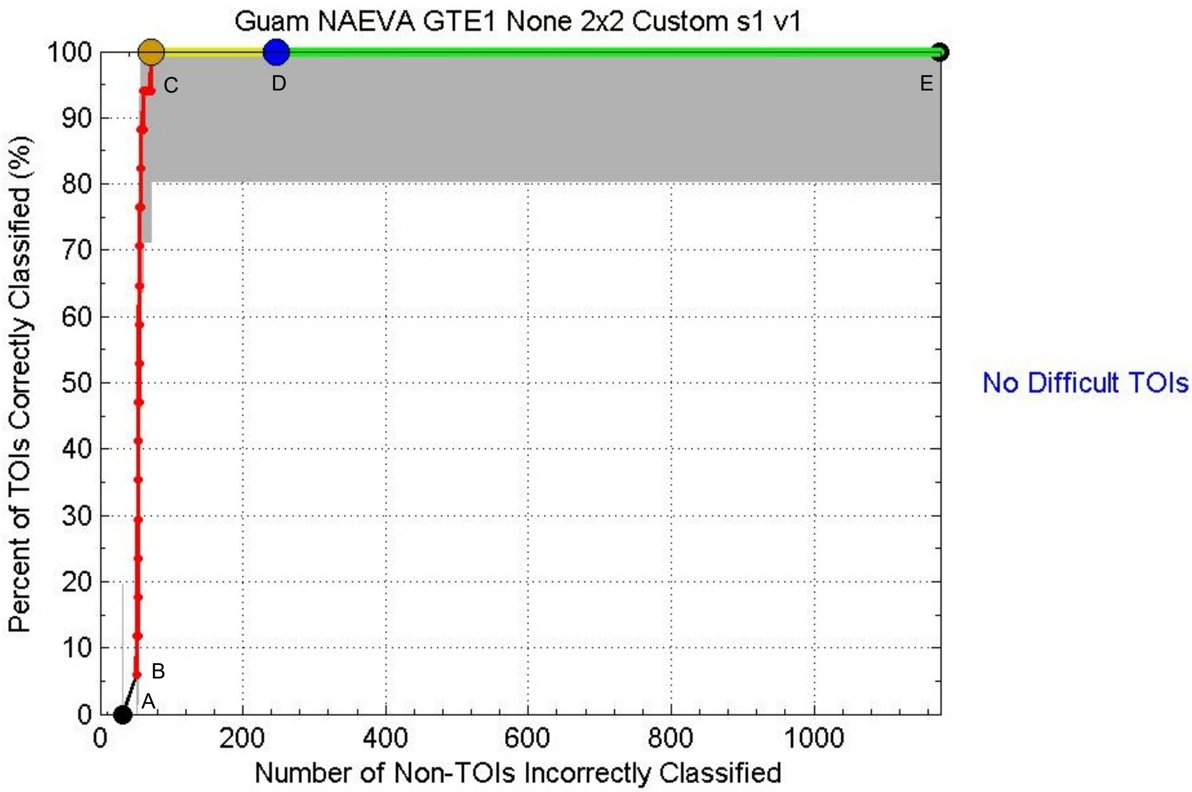

Figure 7-4. GTE1 ROC Curve

Recovered source diameters were not provided with the excavation results. Three of the seed items were explicitly identified, and an additional three specified a general size (i.e., small, medium, large). Of these six targets, 100% were assigned to the correct size group. 18.5% of the estimated source locations were within 15cm of the recovered locations. The accuracy of the recovered source locations does not support the current threshold. No consistent offset was observed between the predicted and recovered locations. There is no available documentation on dig precision, and at the present it is unclear whether the offsets are a result of collection procedures, the processing method, or whether the excavation locations were inaccurately recorded.

All of the TOI items were identified within the top 38 ranked targets, all of which were categorized as “high confidence TOI”. The conservative dig threshold, based largely on the repeatability of the daily IVS measurements, introduced 195 additional non-TOI digs, however the success criteria metric for the reduction of non-TOI digs was sufficiently met with this approach.

7.2.2 GTE2

All seed items were correctly classified as high confidence TOI. No native items were present within the demonstration area. 60 Non-TOI items remained above the dig line, resulting in 1,117 Non-TOI that did not have to be dug. Classification results are summarized in figure 7-5.

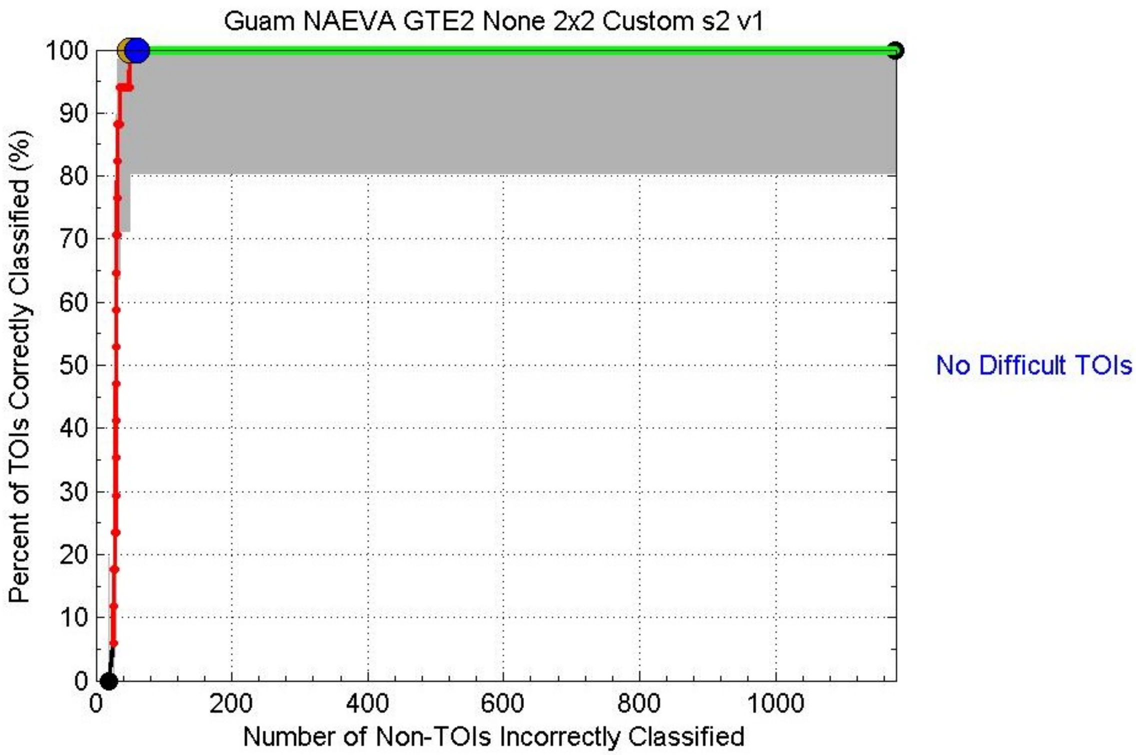

Figure 7-5. GTE2 ROC Curve

Recovered source diameters were not provided with the excavation results. Three of the seed items were explicitly identified, and an additional three specified a general size (i.e., small, medium, large). Of these six targets, 100% were assigned to the correct size group. As seen in the other analysis methods, only a small number of source locations were within 15cm of the recovered locations; 17% for the GTE2 analysis. The source of this error is not known without more information on excavation procedures.

All of the TOI items were identified within the top 67 ranked targets, all of which were categorized as "high confidence TOI". The last identified TOI item was a secondary target pick from lower resolution EM61-MK2 data. The last ranked single TOI was in the top 50 targets. A fairly aggressive dig threshold was chosen based on items selected for training data which resembled small TOI and were identified as non-TOI in the ground truth. The "cannot decide" category was eliminated for the final dig list submittal based on confidence in the classification approach and visual review of each target.

7.2.3 GTE3

All seed items were correctly classified as high confidence TOI. No native items were present within the demonstration area. 365 Non-TOI items remained above the dig line, resulting in 812 Non-TOI that did not have to be dug. Classification results are summarized in figure 7-6.

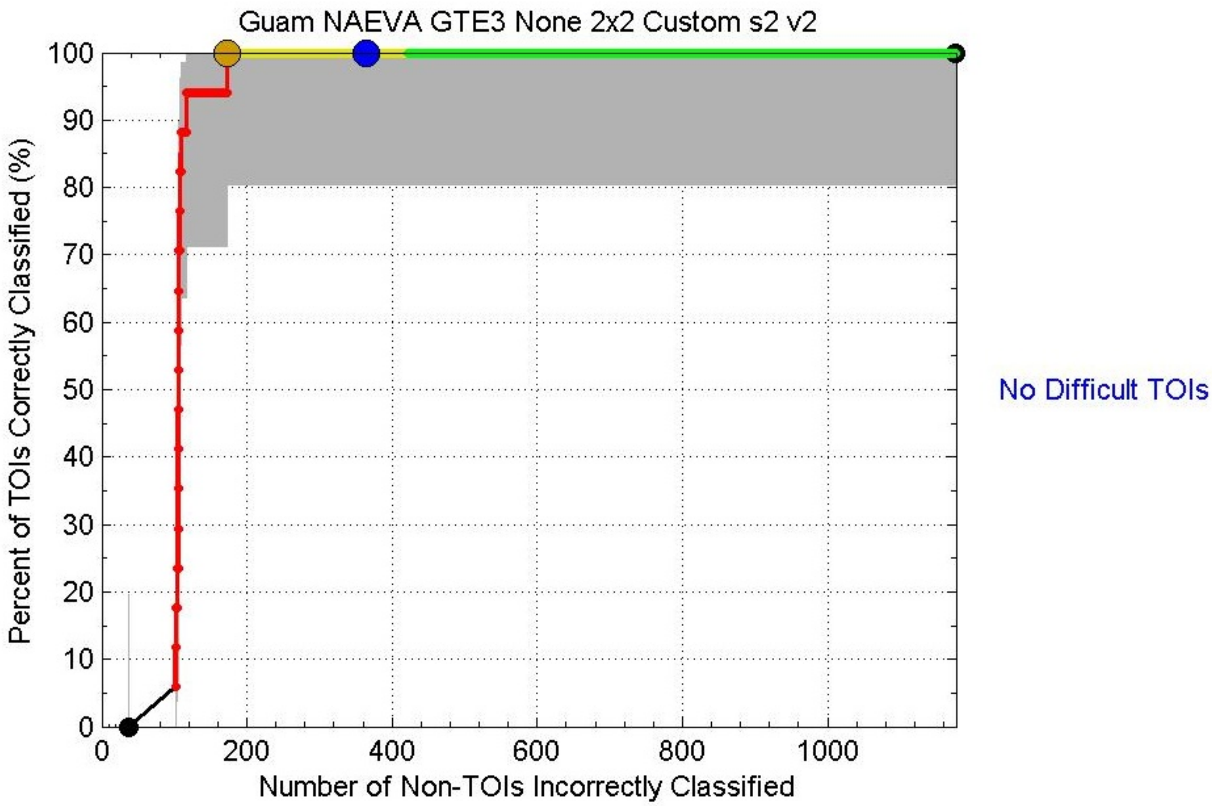

Figure 7-6. GTE3 ROC Curve

Recovered source diameters were not provided with the excavation results. Three of the seed items were explicitly identified, and an additional three specified a general size (i.e. small, medium, large). Of these six targets, 100% were assigned to the correct size group. As seen in the other analysis methods, only a small number of source locations were within 15cm of the recovered locations; 17% for the GTE3 analysis. The source of this error is not known without more information on excavation procedures.

All of the TOI items were identified within the top 87 ranked targets which were categorized as "high confidence TOI". The lowest ranked TOI, target 13341, was manually moved into the "high confidence TOI category by the analyst from the "cannot analyze category". The single and multi-solution solvers indicated this target was located outside the instrument footprint. Visual inspection of the dynamic data and comparison of adjacent targets indicated the source location to be viable, and that the library match to TOI should put this target in the "high confidence TOI" category. Figure 7-7 shows the decision plot for target 13341.

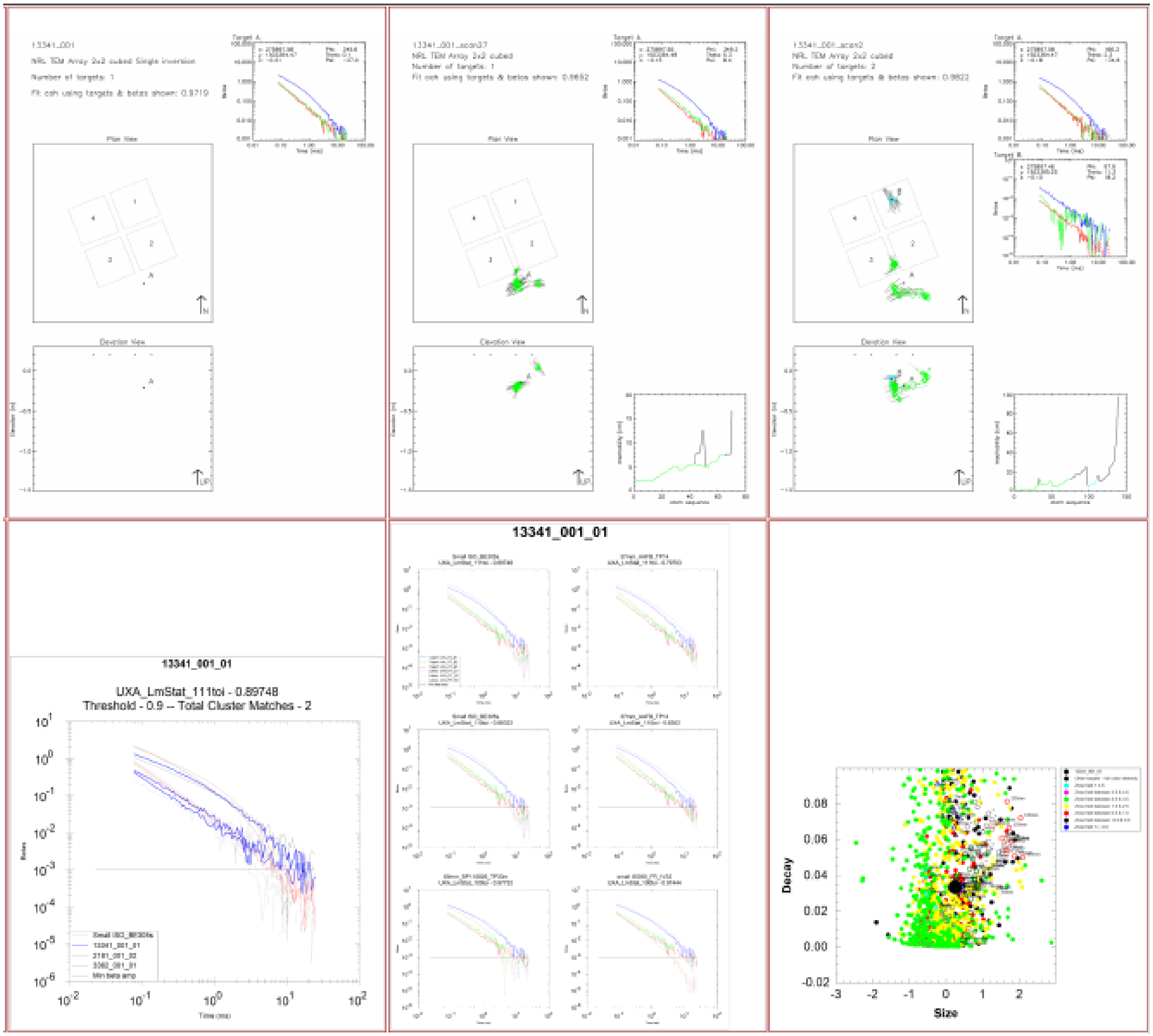

Figure 7-7. Decision Plot for Target GU-13341 Displaying (Clockwise from Upper Left) Single Source Inversion Result, Multisource Inversion Result, multi_110 Source Inversion Result, Size/Decay Plot, Library Match Polarizability Plot to TOI, Library Match Polarizability Plot to TOI with Cluster

The dig threshold was left conservative, introducing 199 additional non-TOI digs, however the success criteria metric for the reduction of non-TOI digs was sufficiently met with this approach. Further selection of training data would have aided in raising the dig threshold by helping to remove some small items that resembled TOI.

7.3 SUMMARY OF PERFORMANCE

NAEVA processors correctly classified all TOI and achieved a substantial reduction in the number of required digs. The limited ground truth available restricts the extent of retrospective analysis possible.

Four of the targets categorized as “High Confidence TOI” in ranked list GTE1have ground truth indicating that cultural debris was found at these locations, but no recovered depths or item descriptions were provided. Removing these items from the scoring results produces a moderate improvement in classification.

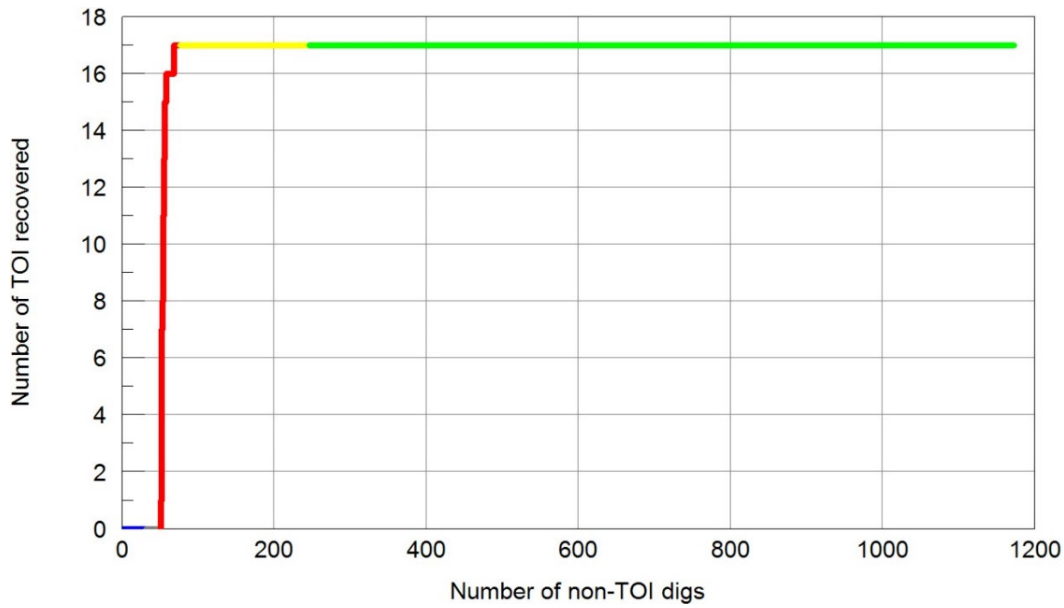

Figure 7-8. GTE1 ROC Curve with Unidentified Clutter Sources Removed

Additionally, the excavation records for some high confidence TOI targets indicate a recovered source that differs substantially from the source indicated by the extracted target parameters. Figure 7-9 shows a decision plot for target GU-781, which was identified during excavation as an unspecified number of nails. With all polarizabilities equally weighted, this target has a confidence metric of 0.018, indicating an incredibly low match, to a target identified from training data as having the same recorded source. The confidence metric for the match between this source and a small ISO80 emplaced in the IVS is over 0.99. This would support the idea that an additional source was present at this flag location and not excavated, or that the source was misidentified in the ground truth.

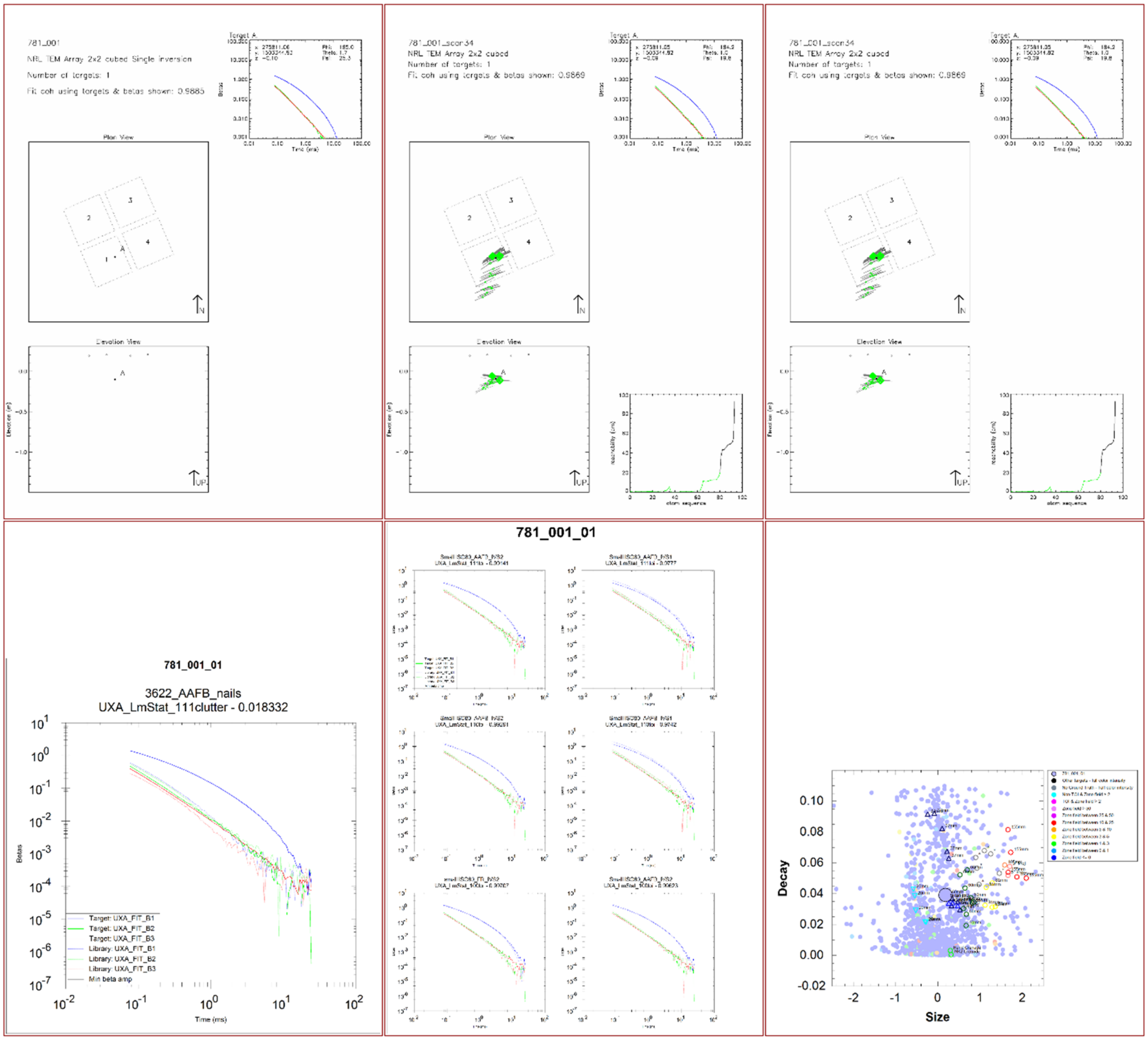

Figure 7-9. Decision Plot for Target GU-781 Displaying (Clockwise from Upper Left) Single Source inversion Result, Multisource Inversion Result, Multi_110 Source Inversion Result, Size/Decay Plot, Library Match Polarizability Plot to TOI, Library Match Polarizability Plot to Nails

8.0 REFERENCES

1. Explosive Safety Submission – Guam Construction Support, Amendment #5, June 2012.

2. Steigerwalt, Ryan. ESTCP Munitions Response Live Site Demonstrations, Anderson Air Force Base, Guam Demonstration Plan, Weston Solutions, Inc. Version 2. 31 December 2013.

APPENDIX A POINTS OF CONTACT

Cued classification was performed by NAEVA Geophysics processors. Contact information is as follows:

NAEVA Geophysics Inc.
P.O. Box 7325
Charlottesville, VA 22906
(434)978-3187 (phone)
(434)973-9791 (fax)

APPENDIX B IVS SUMMARY

UXA_TARGET_ID	UXA_FIT_COH14	UXA_LmStat_111toi_best
1_001_01	0.9860	0.9980
1_002_01	0.9857	1.0000
1_003_01	0.9885	0.9955
1_004_01	0.9822	0.9905
1_005_01	0.9858	0.9934
1_006_01	0.9847	0.9887
1_007_01	0.9876	0.9964
1_008_01	0.9804	0.9939
1_009_01	0.9828	0.9716
1_010_01	0.9873	0.9947
1_011_01	0.9870	0.9954
1_012_01	0.9887	0.9967
1_013_01	0.9879	0.9923
1_014_01	0.9894	0.9908
1_015_01	0.9887	0.9765
1_016_01	0.9884	0.9949
1_017_01	0.9718	0.9926
1_018_01	0.9862	0.9779
1_019_01	0.9866	0.9943
1_020_01	0.9804	0.9914
1_021_01	0.9882	0.9961
1_022_01	0.9895	0.9921
2_001_01	0.9685	0.9978
2_002_01	0.9682	1.0000
2_003_01	0.9704	0.9909
2_004_01	0.9398	0.9724
2_005_01	0.9663	0.9816
2_006_01	0.9594	0.9926
2_007_01	0.9738	0.9762
2_008_01	0.9629	0.9893
2_009_01	0.9321	0.9830
2_010_01	0.9410	0.9772
2_011_01	0.9582	0.9922
2_012_01	0.9772	0.9927
2_013_01	0.9701	0.9867
2_014_01	0.9728	0.9927
2_015_01	0.9671	0.9877
2_016_01	0.9644	0.9906

UXA_TARGET_ID	UXA_FIT_COH14	UXA_LmStat_111toi_best
2_017_01	0.9550	0.9912
2_018_01	0.9713	0.9873
2_019_01	0.9511	0.9789
2_020_01	0.9721	0.9832
2_021_01	0.9691	0.9922
2_022_01	0.9630	0.9931
3_001_01	0.9673	0.9767
3_002_01	0.9705	1.0000
3_003_01	0.9745	0.9516
3_004_01	0.9522	0.9470
3_005_01	0.9838	0.9621
3_006_01	0.9762	0.9423
3_007_01	0.9811	0.9266
3_008_01	0.9653	0.9129
3_009_01	0.9728	0.8432
3_010_01	0.9854	0.9665
3_011_01	0.9793	0.9272
3_012_01	0.9801	0.9617
3_013_01	0.9686	0.9457
3_014_01	0.9822	0.9548
3_015_01	0.9760	0.9505
3_016_01	0.9844	0.9524
3_017_01	0.8071	0.9530
3_018_01	0.9772	0.9698
3_019_01	0.9342	0.9610
3_020_01	0.9766	0.9538
3_021_01	0.9752	0.9576
3_022_01	0.9809	0.9548
4_001_01	0.9632	0.9343
4_002_01	0.9630	1.0000
4_003_01	0.9664	0.9181
4_004_01	0.9539	0.8154
4_005_01	0.9633	0.8712
4_006_01	0.9684	0.8946
4_007_01	0.9692	0.8966
4_008_01	0.9615	0.8711
4_009_01	0.9568	0.7409
4_010_01	0.9925	0.8704
4_011_01	0.9910	0.8525
4_012_01	0.9898	0.9097

UXA_TARGET_ID	UXA_FIT_COH14	UXA_LmStat_111toi_best
4_013_01	0.9894	0.9149
4_014_01	0.9895	0.8657
4_015_01	0.9901	0.7685
4_016_01	0.9929	0.9344
4_017_01	0.9802	0.8130
4_018_01	0.9901	0.8170
4_019_01	0.9654	0.8516
4_020_01	0.9894	0.8084
4_021_01	0.9845	0.8580
4_022_01	0.9924	0.8628
5_001_01	0.9722	0.9675
5_002_01	0.9733	1.0000
5_003_01	0.9757	0.9047
5_004_01	0.9670	0.9127
5_005_01	0.9751	0.9555
5_006_01	0.9774	0.9150
5_007_01	0.9760	0.9283
5_008_01	0.9694	0.8938
5_009_01	0.9724	0.9153
5_010_01	0.9866	0.9403
5_011_01	0.9793	0.8910
5_012_01	0.9866	0.9118
5_013_01	0.9843	0.9000
5_014_01	0.9895	0.9279
5_015_01	0.9907	0.9318
5_016_01	0.9927	0.8741
5_017_01	0.9880	0.9167
5_018_01	0.9881	0.9032
5_019_01	0.9748	0.9314
5_020_01	0.9907	0.8893
5_021_01	0.9857	0.8966
5_022_01	0.9904	0.9035
6_001_01	0.9832	0.7992
6_002_01	0.9842	1.0000
6_003_01	0.9861	0.9072
6_004_01	0.9735	0.8795
6_005_01	0.9807	0.8718
6_006_01	0.9791	0.9005
6_007_01	0.9801	0.9076
6_008_01	0.9745	0.9068

UXA_TARGET_ID	UXA_FIT_COH14	UXA_LmStat_111toi_best
6_009_01	0.9719	0.8432
6_010_01	0.9796	0.8725
6_011_01	0.9779	0.8904
6_012_01	0.9824	0.8961
6_013_01	0.9722	0.8832
6_014_01	0.9804	0.8343
6_015_01	0.9852	0.8611
6_016_01	0.9776	0.8968
6_017_01	0.9806	0.8732
6_018_01	0.9078	0.4254
6_019_01	0.9828	0.8410
6_020_01	0.9661	0.8741
6_021_01	0.9833	0.8886
6_022_01	0.9782	0.8880
6_023_01	0.9825	0.8749
7_001_01	0.9839	0.9845
7_002_01	0.9838	1.0000
7_003_01	0.9657	0.9623
7_004_01	0.9843	0.9819
7_005_01	0.9833	0.9858
7_006_01	0.9834	0.9918
7_007_01	0.9825	0.9810
7_008_01	0.9850	0.9787
7_009_01	0.9823	0.9829
7_010_01	0.9822	0.9799
7_011_01	0.9837	0.9685
7_012_01	0.9837	0.9739
7_013_01	0.9823	0.9876
7_014_01	0.9839	0.9506
7_015_01	0.9867	0.9685
7_016_01	0.9836	0.9680
7_017_01	0.9781	0.9725
7_018_01	0.9870	0.9627
7_019_01	0.9823	0.9892
7_020_01	0.9833	0.9713
7_021_01	0.9833	0.9849
7_022_01	0.9836	0.9751
8_001_01	0.9666	0.9849
8_002_01	0.9661	1.0000
8_003_01	0.9794	0.9875

UXA_TARGET_ID	UXA_FIT_COH14	UXA_LmStat_111toi_best
8_004_01	0.9778	0.9939
8_005_01	0.9785	0.9882
8_006_01	0.9647	0.9895
8_007_01	0.9648	0.9890
8_008_01	0.9659	0.9874
8_009_01	0.9744	0.9905
8_010_01	0.9703	0.9885
8_011_01	0.9774	0.9918
8_012_01	0.9734	0.9893
8_013_01	0.9769	0.9742
8_014_01	0.9718	0.9723
8_015_01	0.9856	0.9843
8_016_01	0.9753	0.9743
8_017_01	0.9648	0.9753
8_018_01	0.9690	0.9770
8_019_01	0.9705	0.9825
8_020_01	0.9655	0.9923
8_021_01	0.9757	0.9871
8_022_01	0.9649	0.9877