Enhanced UXO Discrimination Using Frequency-Domain Electromagnetic Induction

ESTCP Project 200033

Naval Research Laboratory

May 2007

Contents

Figures iv

Tables vii

Abstract A1

1. Introduction 1

1.1 Background 1

1.2 Objectives of the Demonstration 2

2. Technology Description 3

2.1 Technology Development and Applications 3

2.2 Previous Testing of the Technology 5

2.2.1 Array Characterization 5

2.2.2 Array Platform Attitude 7

2.2.3 Motion-induced noise 7

2.2.4 Blossom Point Test Field Survey 9

2.3 Factors Affecting Cost and Performance 10

2.4 Advantages and Limitations of the Technology 10

3. Demonstration Design 11

3.1 Performance Objectives 11

3.2 Selecting the Test Sites 11

3.3 Test Site History/Characteristics 11

3.3.1 Climate and Weather 12

3.3.2 Topography 12

3.3.3 Site Maps and Photographs 12

3.4 Testing and Evaluation Plan 14
3.4.1 Demonstration Set-up and Start-up 14
3.4.2 Period of Operation 14
3.4.3 Area Characterized 20
3.4.4 Operating Parameters of the Technology 21
3.4.5 Experimental Design 21
3.4.6 Demobilization 22
4. Performance Assessment 23
4.1 Aberdeen Proving Ground Blind Grid 23
4.1.1 Overview of the Survey Data 23
4.1.2 Response Stage 29
4.1.3 Discrimination Stage 32
4.2 Aberdeen Proving Ground Open Field 41
4.2.1 Response Stage 41
4.2.2 Discrimination Stage 43
4.3 Yuma Proving Ground Open Field 43
4.3.1 Response Stage 43
4.3.2 Discrimination Stage 45
5. Cost Assessment 46
5.1 Cost Reporting 46
5.2 Cost Analysis 47
6. References 48
7. Points of Contact 49
Appendix A. Summary of the APG Blind Grid Scoring Report 51

Appendix B. Summary of the APG Open Field Scoring Report 54

Appendix C. Summary of the YPG Blind Grid Scoring Report 57

Appendix D. Summary of the YPG Open Field Scoring Report 60

Figures

1. MTADS GEM-3 array shown mounted on the EMI sensor platform with three-member GPS array and IMU shown 3

2. MTADS GEM array surveying at the YPG UXO Demonstration Test Site 4

3. Measured response of the members of the Demonstration array to a calibration ferrite rod located 12.5 cm above the center of each sensor in turn 5

4. Response of the three members of the Demonstration array to a 4 in x 12 in steel cylinder oriented vertically 42 cm above the array 6

5. Illustration of sensor timing 6

6. Measured platform pitch and roll as reported by the GPS array and the IMU 7

7. Comparison of sensor noise at two frequencies for static vs. driving operation of the Demonstration GEM array 8

8. Normalized response (normalized to Q(1230)) to a 16-lb shotput buried at 25 cm 8

9. Measured survey data (in-phase and quadrature response) at three frequencies for the GEM-3 array survey of our Blossom Point test field 9

10. Aerial photograph of the Aberdeen Test Site with the various scenarios outlined 13

11. Aerial photograph of the Yuma Test Site with the various scenarios outlined 14

12. A schematic diagram of the integration of GEM and navigation data for analysis 21

13. Schematic of the sensor platform 22

14. Anomaly image of the 150 Hz in-phase response at the APG Blind Grid 24

15. Anomaly image of the 150 Hz quadrature response at the APG Blind Grid 25

16. Anomaly image of the 11,430 Hz in-phase response at the APG Blind Grid 26

17. Anomaly image of the 11,430 Hz quadrature response at the APG Blind Grid 27

18. Anomaly image of the 11,430 in-phase response at the YPG Blind Grid 28

19. Q_{avg} anomaly image map of the APG Blind Grid 29

20. Detection performance as a function of depth at the APG Blind Grid 30

21. Response stage results showing cumulative ordnance count vs cumulative clutter count 31

22. Response stage performance showing cumulative occupied cell count plotted vs adjusted cumulative blank cell count 31

23. Library response data being acquired for a 60-mm mortar 32

24. Measured response curves for the mortar pictured in Figure 23 32

25. Plot of χ^2 for best match of target responses to library data for two weighting schemes 34

26. The data from Figure 25 plotted as a function of maximum signal amplitude rather than item number 35

27. χ^2 for the Blossom Point Pit data as a function of the amount of "bouncing noise" included in the weighting factor 37

28. Recalculation of the best fit χ^2 for the three sites using a value of 0.3 for the "bouncing noise" coefficient 38

29. χ^2 ratio method applied to the data from the three sites 39

30. ROC curve for the χ^2 weighting applied to the APG Blind Grid 40

31. ROC curve for the case of χ^2 weighting with an estimate of "bouncing noise" included applied to the APG Blind Grid 40

32. ROC curve for the χ^2 ratio method applied to the APG Blind Grid.....41

33. Detection performance at the APG Open Field Scenario.....42

34. Response stage results for the APG Open Field scenario broken out by target type 42

35. Discrimination performance at the APG Open Field Scenario 43

36. Detection performance at the YPG Open Field Scenario.....44

37. Response stage results for the YPG Open Field scenario broken out by target type 44

38. Discrimination performance at the APG Open Field Scenario 45

A1. APG Blind Grid probability of detection for response and discrimination stages versus their respective probability of false positive over all ordnance categories combined 51

A2. APG Blind Grid probability of detection for response and discrimination stages versus their respective probability of background alarm over all ordnance categories combined ..52

B1. APG Open Field probability of detection for response and discrimination stages versus their respective probability of false positive over all ordnance categories combined 54

B2. APG Open Field probability of detection for response and discrimination stages versus their respective probability of background alarm over all ordnance categories combined ..55

C1. YPG Blind Grid probability of detection for response and discrimination stages versus their respective probability of false positive over all ordnance categories combined 57

C2. YPG Blind Grid probability of detection for response and discrimination stages versus their respective probability of background alarm over all ordnance categories combined ..58

D1. YPG Open Field probability of detection for response and discrimination stages versus their respective probability of false positive over all ordnance categories combined 60

D2. YPG Open Field probability of detection for response and discrimination stages versus their respective probability of background alarm over all ordnance categories combined ..61

Tables

1. Performance Objectives for the Demonstration 11

2. First Order Points at the APG Demonstration Site 13

3. First Order Points at the YPG Demonstration Site 13

4. Details of the 2003 Aberdeen Survey 15

5. Details of the 2003 Yuma Survey 16

6. Details of the 2004 Aberdeen Survey 19

7. MTADS GEM array coverage at the two test sites 20

8. Summary of Detection Performance at the APG Blind Grid 30

9. Summary of Costs for a 50-acre GEM array survey 46

10. Measured Productivities of the Three MTADS Sensor Arrays 47

A1. Summary of APG Blind Grid Results 52

A2. APG Blind Grid Efficiency and Rejection Rates 53

A3. Correct Type Classification of Targets Correctly Discriminated As UXO at the APG Blind Grid 53

A4. Mean Location Error and Standard Deviation (m) for the APG Blind Grid 53

B1. Summary of APG Open Field Results 55

B2. APG Open Field Efficiency and Rejection Rates 56

B3. Correct Type Classification of Targets Correctly Discriminated As UXO at the APG Open Field 56

B4. Mean Location Error and Standard Deviation (m) for the APG Open Field 56

C1. Summary of YPG Blind Grid Results 58

C2. YPG Blind Grid Efficiency and Rejection Rates 59

C3. Correct Type Classification of Targets Correctly Discriminated As UXO at the YPG Blind Grid 59

C4. Mean Location Error and Standard Deviation (m) for the YPG Blind Grid 59

D1. Summary of YPG Open Field Results 61

D2. YPG Open Field Efficiency and Rejection Rates 62

D3. Correct Type Classification of Targets Correctly Discriminated As UXO at the YPG Open Field.....62

D4. Mean Location Error and Standard Deviation (m) for the YPG Open Field.....62

Abstract

The Chemistry Division of the Naval Research Laboratory has participated in several programs funded by SERDP and ESTCP whose goal is to enhance the discrimination ability of MTADS. The process is based on making use of both the location information inherent in an item's magnetometry response and the shape and size information inherent in the response to the time-domain electromagnetic induction (EMI) sensors that are part of the baseline MTADS in either a cooperative or joint inversion. In all these efforts, our classification ability has been limited by the information available from the EMI sensor. The EM61 is a time-domain instrument with either a single gate to sample the amplitude of the decaying signal or four gates relatively early in time.

To make further progress on UXO classification, a sensor with more information available is required. The GEM-3 sensor is a frequency-domain sensor with up to ten frequencies available for simultaneous measurement of the in-phase and quadrature response of the target. Thus, in principle, there is much more information available from a GEM-3 sensor for use in classification decisions. Unfortunately, the commercial GEM sensor is a hand-held instrument with relatively slow data rates and is thus not very amenable to rapid, wide area surveys. ESTCP Project 200033, Enhanced UXO Discrimination Using Frequency-Domain Electromagnetic Induction, has been funded to overcome this limitation by integrating an array of GEM sensors with the MTADS platform.

The objective of this program was to demonstrate the optimum system that delivers the most classification performance while retaining acceptable survey efficiency. Working around a modified GEM-3 sensor (modified based on our observations during an initial, sensor-characterization phase of the program), we have designed a three-sensor array and demonstrated it at the Standardized UXO Demonstration sites at Aberdeen Proving Ground and Yuma Proving Ground. At each of the sites, we surveyed the Calibration Lanes, the Blind Test Grid, and as much of the Open Field Area as was possible. For the Blind Test Grid and the Open Field, our ranked target picks were submitted to Aberdeen Test Center for scoring. These scoring results are the basis for judging the success of the demonstrations.

Each of the sensors in the array sequentially transmits a composite waveform made up of nine frequencies logarithmically spaced from 90 Hz to just over 20 kHz for one base period (1/30 s). Thus, only three complete cycles of the 90 Hz frequency is transmitted while many thousands of cycles of the highest frequency are transmitted. The transmit current drives both a transmit coil and a counterwound bucking coil. This serves to set up a "magnetic cavity" inside the bucking coil in which is placed a receive coil. The current induced in this receive coil by the induced fields in buried metal targets is detected, digitized, and frequency resolved during the two subsequent base periods while the other array sensors are transmitting. The detected signal is compared to the transmitted current and reported relative to the transmit current (ppm) as both an in-phase and quadrature component.

Individual sensors in the array are located using a three-receiver RTK GPS system. From this set of receivers, we record the position of the master antenna at 20 Hz, and the vectors to the other two antennae at 10 Hz. All positions are recorded at full RTK precision, ~2-5 cm. In addition, we record the output of a full 6-axis IMU at 80 Hz to give complimentary information on platform pitch and roll. All sensor readings are referenced to the GPS 1-PPS output so we are able to fully take advantage of the precision of the GPS measurements.

In this report, we summarize our initial testing of the commercial GEM-3 sensor and the integration and shake-down testing of the array of modified sensors. We then present our Demonstration results focusing on the Aberdeen Proving Ground Blind Grid to describe our detection and discrimination methodologies and performance. We present the results of our initial analysis scheme and two others that we have since developed. We then present summaries of our Open Field scenario performance at the two sites based on extracted results provided to us by analysts at the Institute for Defense Analyses. Finally, in the Appendices we present the key results from the official scoring reports for the various Demonstration surveys.

Enhanced UXO Discrimination Using Frequency-Domain Electromagnetic Induction

1. Introduction

1.1 Background

Unexploded ordnance (UXO) detection and remediation is a high priority triservice requirement. As the Defense Science Board recently wrote: "Today's UXO cleanup problem is massive in scale with some 10 million acres of land involved. Estimated cleanup costs are uncertain but are clearly tens of billions of dollars. This cost is driven by the digging of holes in which no UXOs are present. The instruments used to detect UXOs (generally located underground) produce many false alarms, - i.e., detections from scrap metal or other foreign or natural objects -, for every detection of a real unexploded munition found" [1].

There has been considerable progress in the detection of buried UXO in the last ten years. The Multi-sensor Towed Array Detection System (MTADS), supported by the Environmental Security Technology Certification Program (ESTCP), has demonstrated detection capability for all military ordnance to its maximum self-burial depth [2] with location accuracies [3] on the order of 15 cm. Discrimination of UXO from ordnance fragments and other clutter remains as a problem, however. We have shown that with careful mission planning and a modest on-site training effort, an MTADS survey/remediation project can achieve false alarm rates substantially lower than those quoted for "mag and flag" surveys. However, there is still much room for improvement in discrimination ability that will result in direct reduction of remediation costs.

The Chemistry Division of the Naval Research Laboratory has participated in several programs funded by SERDP and ESTCP whose goal is to enhance the discrimination ability of MTADS. The process is based on making use of both the location information inherent in an item's magnetometry response and the shape and size information inherent in the response to the time-domain electromagnetic induction sensors that are part of the baseline MTADS in either a cooperative or joint inversion. We have already made significant progress toward our goal. The algorithms and methods that were developed for the ESTCP program involving analysis of data from the MTADS EM61 array were applied in preliminary form at the JPG-IV Demonstration allowing us to score as one of the small group of approaches that showed any classification ability [4]. More recently, we have used the methods at JPG-V and on a live range, the Impact Area of the Badlands Bombing Range, SD. In all these demonstrations, our classification ability has been limited by the information available from the sensor. The EM61 is a time-domain instrument with a single gate to sample the amplitude of the decaying signal. To make further progress on UXO classification, a sensor with more information available is required.

Manuscript approved XX June 2007

By far the best results at JPG-IV were obtained by Geophex and AETC using magnetometers and the GEM-3 frequency-domain electromagnetic induction sensor [4]. The GEM-3 sensor is a frequency-domain sensor with up to ten frequencies available for simultaneous measurement of the in-phase and quadrature response of the target. Thus, in principle, there is much more information available from a GEM-3 sensor for use in classification decisions. Unfortunately, the commercial GEM sensor is a hand-held instrument with relatively slow data rates and is thus not very amenable to rapid, wide area surveys. ESTCP Project 200033, Enhanced UXO Discrimination Using Frequency-Domain Electromagnetic Induction, has been funded to overcome this limitation by integrating an array of GEM sensors with the MTADS platform.

1.2 Objectives of the Demonstration

The original objective of the demonstration was simple; demonstrate the classification performance of the GEM-3 sensor in conjunction with the survey efficiency of the MTADS. As we undertook the first task of the program, specification of the sensor configuration most amenable to platform-based array use, we quickly became aware of a significant noise issue with the original GEM-3 sensor when mounted on a cart. Our observations and subsequent modifications to the original sensor are detailed in an earlier report [5].

The revised objective is only slightly more complicated; demonstrate the optimum system that delivers the most classification performance while retaining acceptable survey efficiency. Working around a modified GEM-3 sensor, we have designed a three-sensor array that will be described in detail in the next section and demonstrated it at the Standardized UXO Demonstration sites at Aberdeen Proving Ground and Yuma Proving Ground. At each of the sites, we surveyed the Calibration Lanes, the Blind Test Grid, and as much of the Open Field Area as was possible. For the Blind Test Grid and the Open Field, our ranked target picks were submitted to Aberdeen Test Center for scoring. These scoring results are the basis for judging the success of the demonstrations.

2. Technology Description

2.1 Technology Development and Applications

Based on the lessons learned with the prototype array [5], we arrived at a design for the demonstration array which is shown in Figure 1. The coils used in the array are 96-cm diameter with a larger number of turns. New, higher current electronics have been designed and constructed to drive the coils. The product of these two factors results in a factor of six to eight increase in transmit moment. The new sensor electronics bring other benefits as well. We are now able to implement real-time low-pass filtering of the induced current using a DSP chip in the receive circuit. This reduces the absolute level of the noise in addition to the relative decrease from the increased transmit moment. Other deployment measures have been taken to reduce the induced noise in the sensor array. The three sensors are mounted on a rigid base attached to the sensor tray to minimize relative motion. We have optimized the sensor cart for steady ride. This includes making the wheelbase as long as possible and controlling tire pressure. We also deploy an array of GPS receivers to measure orientation of the sensor platform as well as position and a six-axis inertial measurement unit (IMU).

Fig. 1 – MTADS GEM-3 array shown mounted on the EMI sensor platform with three-member GPS array and IMU shown

The sensor timing sequence was chosen to be non-synchronous (each sensor operating independently) and sequential. This was done to avoid the expense of developing new drive electronics (the individual GEM-3 sensors are configured to run independently) and the large, unbucked inter-sensor interference that results from simultaneous operation. Inter-sensor sequencing is controlled by the custom electronics package mentioned above that is rack-mounted in the MTADS tow vehicle, shown in Figure 2.

Fig. 2 – MTADS GEM array surveying at the YPG UXO Demonstration Test Site

Sequential operation of the sensors does not cause the large reduction in array sampling rate that might be imagined on first thought. The baseline GEM-3 sensor operates on a 1/30 second base period. That is, each sensor acquires data for a multiple of this base period. It then takes approximately two base periods to perform the internal data reduction calculations before the results of the measurement are available to the recording computer. If each of the three sensors in this sequential array are configured to record for a single base period, they can then be run near their maximum rate. The effective sampling rate for the three sensor array is ~9.5 Hz which coupled with a survey speed of ~3 mph results in a down-track sampling interval of ~15 cm. This results in an acceptable data density for our analysis purposes.

Each of the sensors in the array sequentially transmits a composite waveform made up of nine frequencies logarithmically spaced from 90 Hz to just over 20 kHz for one base period (1/30 s) as mentioned above. Thus, only three complete cycles of the 90 Hz frequency is transmitted while many thousands of cycles of the highest frequency are transmitted. The transmit current drives both a transmit coil and a counterwound bucking coil. This serves to set up a "magnetic cavity" inside the bucking coil in which is placed a receive coil. The current induced in this receive coil by the induced fields in buried metal targets is detected, digitized, and frequency resolved during the two subsequent base periods while the other array sensors are transmitting. The detected signal is compared to the transmitted current and reported relative to the transmit current (ppm) as both an in-phase and quadrature component.

Individual sensors in the array are located using a three-receiver RTK GPS system as shown in Figure 1. From this set of receivers, we record the position of the master antenna at 20 Hz, and the vectors to the other two antennae at 10 Hz. All positions are recorded at full RTK precision, ~2-5 cm. In addition, we record the output of a full 6-axis IMU at 80 Hz to give complimentary information on platform pitch and roll. All sensor readings are referenced to the GPS 1-PPS output so we are able to fully take advantage of the precision of the GPS measurements.

The individual data streams into the data acquisition computer, which is rack-mounted in the Tow Vehicle and running a custom variant of the Geophex WinGEM program called WinGEMArray, are each recorded in a separate file. These individual data files, which share a root name that corresponds to the data and time the survey was initiated, include three sensor data files, four GPS files (one containing the NMEA GGK sentences corresponding to the

position of the master antenna and an AVR sentence giving one of the vectors to the secondary antennas, another containing the second AVR sentence, a third containing the UTC time tag, and the fourth containing the computer-time stamped arrival of the GPS PPS), and one file for the IMU output. The sensor and GPS files are ASCII format and the IMU file mirrors the packed binary output of the IMU.

All these files are transferred to the Data Analysis System using ZIP-250 disks. It is here that the data is checked for quality and the files are consolidated to give a product ready for analysis work – i.e. calibrated and leveled sensor readings (in-phase and quadrature response at nine frequencies) correlated with the sensor location and orientation information. The details of the consolidation process are described in section 3.4.5.

2.2 Previous Testing of the Technology

2.2.1 Array Characterization

Our first step in preparing for the demonstrations was to perform a series of static characterization measurements on the array. The first of these was to measure the response of the array to a calibration ferrite rod. The results of this measurement are shown in Figure 3. Each of the curves in the Figure is the average of 5 individual measurements. The spread in the individual measurements is scarcely larger than the width of the lines. As can be seen, sensors 1 and 3 have very similar responses with sensor 2 slightly different. These relative responses are stored and used to normalize all further frequency-dependent measurements.

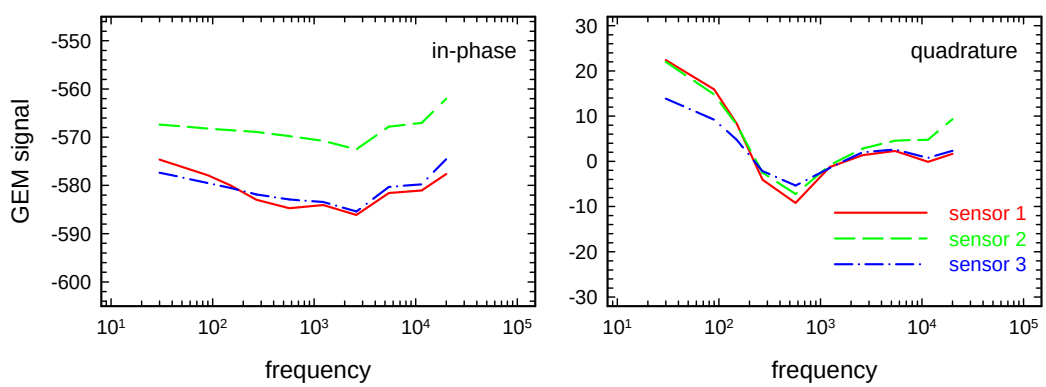

Fig. 3 – Measured response of the members of the Demonstration array to a calibration ferrite rod located 12.5 cm above the center of each sensor in turn

With this calibration in hand, we measured the response of the three members of the array to a set of ordnance simulants. An example of the results of these measurements is shown in Figure 4. These data result from the translation of a 3-in diameter, 12-in long steel cylinder, designated target E, in two orthogonal passes over the array with the target oriented vertically, 42 cm above the array. These are, of course, very large signals so the SNR is quite large. Nevertheless, we see that the response of the three sensors is approximately equal and of the correct form. These data can be inverted to frequency-dependent response coefficients, β, and sensible values result.

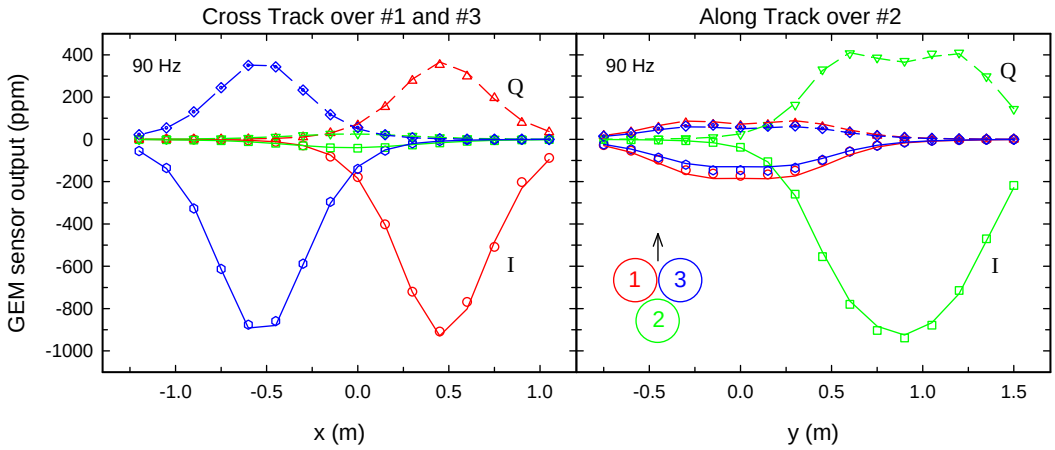

Fig. 4 – Response of the three members of the Demonstration array to a 4 in x 12 in steel cylinder oriented vertically 42 cm above the array

The final issue in characterization of the array relates to absolute and inter-sensor timing. The MTADS is designed with the idea that all hardware latencies will be measured and corrected for. This implies that no survey data set should exhibit any apparent latency. To measure this, we typically drive a number of back-and-forth passes over a metal pipe planted across the top of the test field at Blossom Point. An example of the data collected on three such back-and-forth passes is shown in Figure 5 which shows, for each member of the array, the path taken over the calibration pipe (upper panel) and the measured quadrature signal at 1230 Hz vs. the measured y position (lower panel). Since sensors 1 and 3 are off the centerline, they switch position depending on whether the array is being drive N to S or S to N. The positions plotted in the three lower panels include a constant 37.5 ms sensor latency.

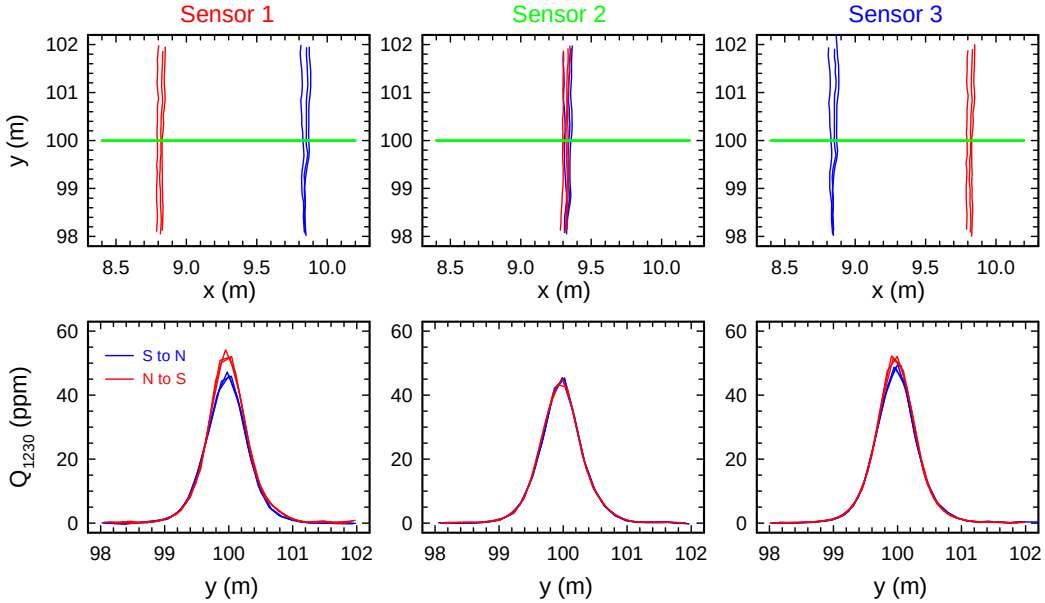

Fig. 5 – Illustration of sensor timing. See the text for an explanation of the plots.

2.2.2 Array Platform Attitude

As previously noted, platform orientation is measured by both the GPS array and the IMU. A comparison of measurements from the two systems is shown in Figure 6. As can be seen, the GPS and IMU results track well with two differences. Obviously, the IMU captures higher frequency information than the 10 Hz GPS measurement. Unfortunately, the IMU suffers drift on timescales much shorter than a typical survey. This is shown in Figure 6 by the offset between the GPS and IMU results. To determine platform orientation during a survey, we combine the two measurements as shown by the third trace in Figure 6 to capture both the accuracy of the GPS and the precision of the IMU.

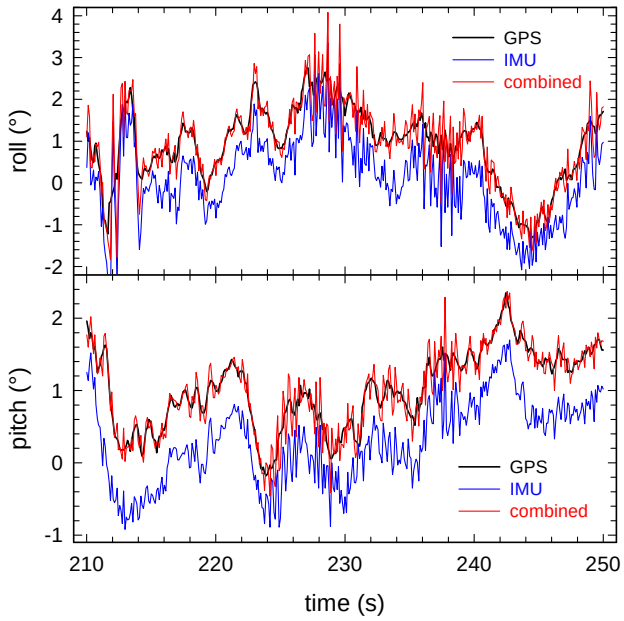

Figure 6 – Measured platform pitch and roll as reported by the GPS array and the IMU. In addition, the combined result which is used in our analysis is shown.

2.2.3 Motion-induced noise

The final component of array characterization is a measurement of motion-induced system noise under realistic deployment conditions. Measurements taken on the test field at Blossom point are plotted in Figure 7. The difference from the early data discussed in [5] is dramatic. The low-frequency noise is greatly reduced. Notice that there is one target detected in the 5430 Hz data. The remaining in-phase noise at higher frequencies results from moving the coil closer and further from the ground as the sensor bounces over the test plot. Unlike the situation at low frequencies, this noise is observed to be coherent across all frequencies (i.e. profile data for all of the frequencies track each other within a percent or two when no metal target is present).

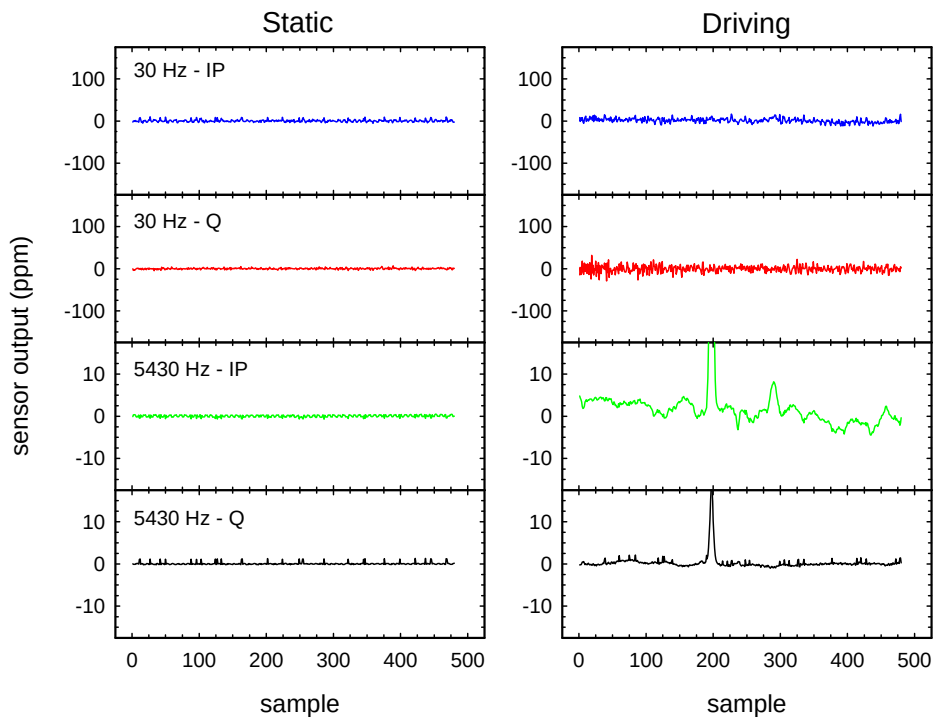

Figure 7 – Comparison of sensor noise at two frequencies for static vs. driving operation of the Demonstration GEM array

Another illustration of the reduction of noise achieved in the Demonstration array is found by examining the response of the sensors to one of the targets in the test field. As the array traverses the test field during a survey, each of the sensors makes a number of passes in the vicinity of each target. Figure 8 shows the normalized responses to a 16-lb shotput buried 25 cm below the surface. Each measured response has been normalized to the quadrature response at 1230 Hz. The 42 response curves are further divided into those with Q(1230) > 30 ppm (16 measurements) which we refer to as high SNR and those with Q(1230) between 5 and 30 ppm (26 measurements).

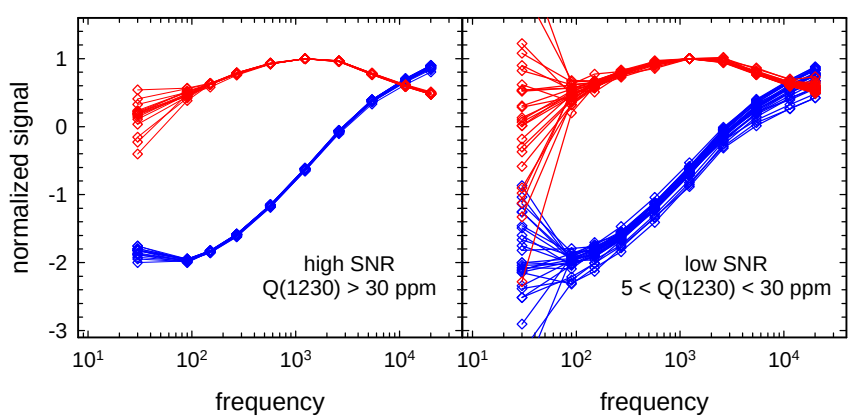

Fig. 8 – Normalized measured response (normalized to Q(1230)) to a 16-lb shotput buried at 25 cm

As can be seen in the left panel of Figure 8, the signal (with the exception of the quadrature response at 30 Hz) is remarkably reproducible for measurements with Q(1230) ranging from 170 ppm to 30 ppm. This is the reproducibility and precision required for successful classification of targets. The spectra shown in the lower panel of Figure 8, while sufficient for detection, would make classification much more difficult.

2.2.4 Blossom Point Test Field Survey

The initial deployment of the GEM-3 array was on our test field at Blossom Point [6]. A plot of the measured response of the system at three frequencies (in-phase and quadrature) is shown in Figure 9. Several points can be derived from these data. The line of targets at x = -6m consists of sets of three 60mm mortars buried at (from y = 69 down) 0. 5m, 0.25m, 1.0m and 0.75m. As can be seen from the figure, the mortars at 0.25 and 0.5 m are easily detectable. The mortars at 0.75m are detectable in an extremely clean field such as this one but would likely not be detectable in a field with a normal distribution of fragments and clutter. The 60mm mortars at 1.0 m are not detectable. The pipe used for timing tests is visible at y = 100m as is a large soil anomaly shown in the 20010 Hz quadrature response.

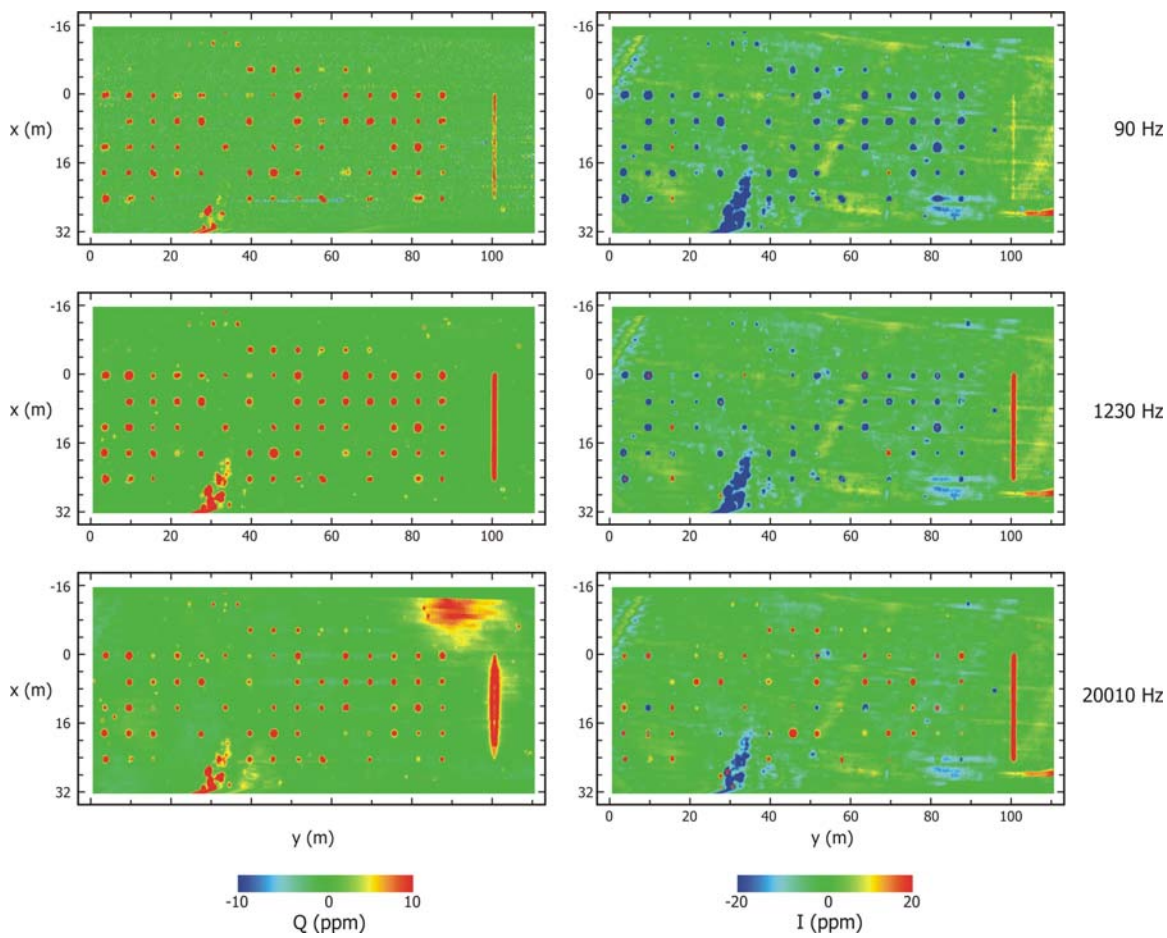

Fig. 9 – Measured survey data (in-phase and quadrature response) at three frequencies for a GEM-3 array survey of our Blossom Point test field

2.3 Factors Affecting Cost and Performance

Implementation of the methods in this Demonstration requires additional survey time compared to a minimum detection survey. We have shown that, in many cases, the MTADS can detect essentially all UXO with a total-field magnetometry survey. For ordnance target sets that include 60- and 81-mm mortars at depths of 0.75 to 1 m and/or 20- and 30-mm submunitions an overlapping EM induction survey is required to get a high detection probability. Compared to the magnetometer array, we use two interleaved GEM array surveys to ensure sufficient data density over each target to get a reliable fit to our models. This increases the survey hours on-site although it does not impact the mobilization and data analysis costs. In many cases, the extra survey costs are only equivalent to the cost of digging one or two additional targets per acre.

2.4 Advantages and Limitations of the Technology

No single method currently available is the "magic bullet" of classification. We have previously demonstrated [3] that some discrimination is possible using the standard MTADS if a small training area is investigated prior to data analysis on the entire site and the distribution of ordnance is limited. This discrimination is based primarily on fitted dipole "size." More recently we have shown that increased discrimination can be achieved by adding an extra "dimension" to the discrimination; that of "shape." For items of the same induced magnetic dipole we can discriminate based on the ratio of responses of the items' three axes to the EM induction sensors in the MTADS suite. As we have shown [7], this adds some discrimination capability to the system.

Unfortunately, the EM methods previously demonstrated are not sufficiently robust to the sensor noise, position uncertainty, and platform motion inherent at a real site. We have proposed that some of this observed drop in performance at real sites can be recaptured using a sensor with more inherent information content. The 9-frequency GEM sensor that is the subject of this demonstration is such a sensor. The depth performance of the GEM sensors is limited compared to the time-domain EM sesnors we have demonstrated in previous programs. This will be expanded upon in the discussion of our results from the Aberdeen test site.

Even with the most optimistic result however, these methods will not result in a perfect system. As we have stated many times, this program is based on the idea of classification by shape. By definition, this implies that clutter items that have similar shapes to ordnance will be classified as ordnance. Items such as pipes and post sections are representative of this problem. If it is important to reduce remediation costs to the extent that these items are not dug, other methods, possibly sensitive to composition or the presence of explosive compounds, will have to be employed in conjunction with those being demonstrated in this program.

3. Demonstration Design

3.1 Performance Objectives

There were two primary objectives of this Demonstration. First, to demonstrate the MTADS GEM array on ranges away from our home location at Blossom Point. This provides a measure of system reliability and ease of use "on the road." Second, to evaluate the probability of detection and classification ability of the combined hardware/analysis system in a blind test. Details of our performance criteria and the expected performance are given in Table 1.

Table 1. Performance Objectives for the Demonstration.

Type of Performance Objective	Primary Performance Criterion	Expected Performance
Qualitative	1. Reliability	Routine survey performance 6 hr/day or more.
	1. Ease of Use	A four-person team is able to perform all data collection and analysis tasks.
	2. Maintenance	All routine maintenance is accomplished on site and all required spare parts are in the inventory.
Quantitative	1. Probability of Detection	
	2. False Alarm Rate	
	3. Array SNR	

3.2 Selecting the Test Sites

The location of the test sites were selected by the PI of the Standardized UXO Technology Test Site Program. We demonstrated this technology at both the Aberdeen Proving Ground and Yuma Proving Ground sites.

3.3 Test Site History/Characteristics

The area of the Aberdeen Test Site is adjacent to the Trench Warfare facility at the Aberdeen Proving Ground. The specific site was obviously used for a variety of ordnance tests over the years, our initial magnetometer and EMI survey, performed after a "mag and flag" survey of the area identified over a thousand remaining anomalies. These data were used for a final clean up of the site prior to test target emplacement.

The Yuma Test Site is adjacent to the Mine Test Area at YPG. The area was reasonably clear before establishment of the Standardized UXO Technology Site; our initial survey found far fewer residual anomalies after the initial clearance than was the case for the Aberdeen site.

3.3.1 Climate and Weather

During the last week in September and the first week in October, the normal high for Aberdeen is in the low 70s F and the normal low is in the low 50s. Average precipitation is 0.12" each day. In 2003, temperatures were normal but precipitation was more than twice the long-term average. In fact, there were two tropical events in the seven days before our survey. This rendered much of the test site too wet for a survey. We returned to Aberdeen in June 2004, to complete the survey. During this survey temperatures were seasonal and precipitation was low.

The Yuma Demonstration was performed in November 2003. During the month of November, the normal high is in the low 80s F and the normal low is in the high 30s. Average precipitation is 0.2" for the month. The average winds are 4 mph. During the period of our survey, temperatures were normal but we again experienced an unusual rain event. This caused a one-day delay in starting the survey but did not affect our performance otherwise.

3.3.2 Topography

The portions of the Aberdeen site covered in this Demonstration (the Calibration Lanes, Blind Test Grid, and Open Field) are relatively flat and level. There are some low-lying areas in the NW portion of the Open Field area that tend to have standing water during wet periods of the year. We were able to survey through this standing water with the MTADS magnetometer system but would not have been able to take the GEM array through water of such depth. As it turned out, the standing water was much worse during our first visit. In June 2004, the water was again at normal levels.

The scenarios at the Yuma site that were covered in this Demonstration (the Calibration Lanes, Blind Test Grid, and Open Field) are also relatively flat and level. Due to the presence of washes and scattered trees on the Open Field site, we were not able to cover every square meter of the site with the vehicular towed system. We surveyed along the edge of the washes on both sides and then down in them as was feasible. The coverage we achieved in the Open Field site was representative of vehicle-towed systems.

3.3.3 Site Maps and Photographs

An aerial photograph of the Aberdeen test site is shown in Figure 10. The various scenarios are outlined and numbered. The scenarios of concern here are the calibration lanes (1) and blind test grid (2) outlined in blue and the open field (3) outlined in yellow. We did not survey either the mogul (4) or wooded (5) areas.

Fig. 10 – Aerial photograph of the Aberdeen Test Site with the various scenarios outlined

There are two first-order points on the site for use as GPS base station points. Their reported coordinates are listed in Table 2. All surveys were conducted using monument 477.

Table 2. First Order Points at the APG Demonstration Site.

Monument	Latitude	Longitude	HAE (m)
477	39º 28' 18.63880" N	76º 07' 47.71815"W	-22.545
478	39º 28' 04.24219" N	76º 07' 48.50439"W	-21.473

A corresponding aerial photograph of the Yuma test site is shown in Figure 11. The various scenarios are outlined. The scenarios of concern here are the calibration lanes and blind test grid and the open field. We did not survey either the mogul or Yuma Desert Brush areas. The available monuments at YPG are listed in Table 3. All surveys were conducted using UXO4.

Table 3. First Order Points at the YPG Demonstration Site.

Monument	Northing (m)*	Easting (m)*	HAE (m)
UXO1	3638488.216	757364.509	133.175
UXO2	3638127.066	757482.139	131.835
UXO3	3638127.910	757656.648	132.848
UXO4	3638338.802	757653.745	134.019

* UTM Zone 11N

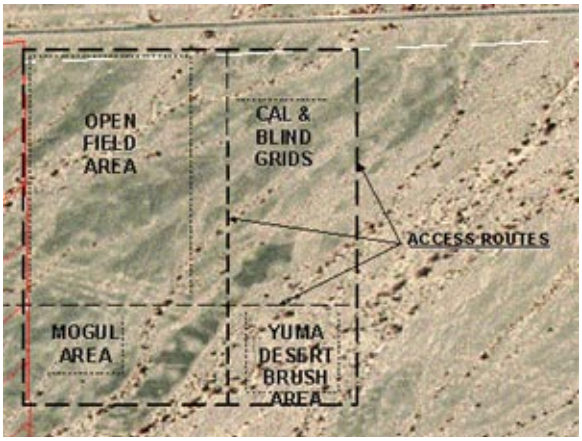

Figure 11 – Aerial photograph of the Yuma Test Site with the various scenarios outlined

3.4 Testing and Evaluation Plan

3.4.1 Demonstration Set-up and Start-up

The MTADS tow vehicle and GEM array trailer were mobilized to the Aberdeen test site on a pair of roll-back trucks. Tools, spares, and miscellaneous supplies were transported in a rented panel van. Since the survey site is only a two and a half hour drive from our base at Blossom Point we only mobilized items required for routine maintenance and immediate field use. As is more usual for an MTADS survey, the tow vehicle, trailer and our complete suite of equipment and spares were mobilized to the Yuma site in a rented 53' trailer equipped with custom mounts for the vehicle and trailer. The trailer was transported to and from Yuma by a government-contract hauling firm.

Since the UXO Technology Demonstration Test Sites are well equipped for surveying, there were no site preparation activities required at either site. Set-up only involves unpacking and assembling the equipment, setting up the GPS base station on a survey monument, warming up the sensors, and collecting a small test data set to verify valid operation of the sensors. At both sites we were able to collect the first real survey data shortly after noon on the first day.

3.4.2 Period of Operation

The Aberdeen Demonstration was scheduled to begin on Monday, Sept. 22, 2003. There was a tropical event in the area the week before and hard rain that day. In consultation with the site managers, we pushed the start date back to Wednesday, Sept 23. The test site was heavily scheduled at the time so we were not able to extend past the two-week window we were originally allotted. Even by Wednesday, the site was too wet for a complete survey. We surveyed the Calibration Lanes and Blind Grid (although we had to skip the SW corner of the Blind Grid) and started the Open Field. After getting stuck several times, we were advised by the site manager that the field was too wet to allow continued surveying. We left the site mid-afternoon Thursday, Sept. 24. After allowing a week for the site to dry, we returned on

Thursday, October 2 to collect a complete survey of the Blind Grid. Details of this survey are given in Table 4.

Table 4. Details of the 2003 Aberdeen Survey.

Date	Description	Survey Base Name	Duration (min)
Wed, Sep 24th	Ferrite Calibration	Sep242003_152428	4
	Warm Up	Sep242003_153728	6
	Timing Calibration	Sep242003_152934	5
	NS Survey of Calibration & Blind Grids	Sep242003_155028 Sep242003_161558 Sep242003_163636 Sep242003_171859	16 8 38 4
	EW Survey of Calibration Grid	Sep242003_175024	22
	EW Survey of Blind Grid (includes area between Calibration & Blind Grids)	Sep242003_181521 Sep242003_182032 Sep242003_185825	5 33 12
	Start NS Survey of Open Field	Sep242003_194537	21
Thur, Sep 25th	Warm Up	Sep252003_123348	6
	Ferrite Calibration	Sep252003_124042	3
	Timing Calibration	Sep252003_124439	3
	Continue NS Survey of Open Field (stopped when no longer feasible to survey because of standing water)	Sep252003_125106 Sep252003_130511	13 46
	Timing Calibration	Sep252003_153716	3
	Ferrite Calibration	Sep252003_154238	9
	Try EW Survey of (dry section of) Open Field	Sep252003_160923 Sep252003_172307	23 46
Thur, Oct 2nd	Warm Up	Oct022003_131552	10
	Ferrite Calibration	Oct022003_132542	4
	Timing Calibration	Oct022003_133841	3
	NS Survey of Calibration & Blind Grids (interleaved tracks to obtain 0.25-m lane spacing)	Oct022003_134512 Oct022003_150356	48 16
	Timing Calibration	Oct022003_152134	3
Total Survey Time			**351**
Total Survey Plus Calibration Time			**388**

The Aberdeen site was heavily scheduled in late 2003 so the next deployment of the MTADS GEM array was to the Yuma Proving Ground site in November. We arrived at the site on Wednesday, November 12th and unpacked and assembled the equipment. Just as in the case at Aberdeen, Yuma experienced an unprecedented rain that afternoon causing the site to be closed for the day. Fortunately, the desert dries much faster than Aberdeen so we were able to begin the survey on Thursday, November 13th, having only lost one day to the weather. Complete details of the survey are given in Table 5.

Table 5. Details of the 2003 Yuma Survey.

Date	Description	Survey Base Name	Duration (min)
Thurs. Nov 13th	Warm Up	Nov132003_134107	12
	Ferrite Calibration	Nov132003_135249	4
	Timing Calibration	Nov132003_135806	3
	4" Al Sphere Calibration	Nov132003_140230	8
	Survey Calibration Lanes	Nov132003_141400	43
	Survey Blind Grid	Nov132003_153514	64
	Survey Open Field	Nov132003_165218 Nov132003_172910	33 32
	Ferrite Calibration	Nov132003_203434	4
	4" Al Sphere Calibration	Nov132003_203856	4
	Timing Calibration	Nov132003_204350	3
	Resurvey Lanes Q & T of Blind Grid	Nov132003_205205	10
	Continue Survey of Open Field	Nov132003_211652 Nov132003_222226	61 58
Fri, Nov 14th	Warm Up	Nov142003_140055	13
	Ferrite Calibration	Nov142003_141423	4
	4" Al Sphere Calibration	Nov142003_141854	5
	M75 "Chute" Up @ 17cm	Nov142003_142419	1
	M75 "Chute" Down @ 17cm	Nov142003_142603	1
	M75 Side @ 16cm	Nov142003_142736	1
	60mm Mortar Nose Down @ 27cm	Nov142003_142851	1
	60mm Mortar Nose Up@ 27cm	Nov142003_143038	2
	60mm Mortar Side @ 16cm	Nov142003_143222	2
	M75 "Chute" Up @ 37cm	Nov142003_143429	1

Date	Description	Survey Base Name	Duration (min)
Fri, Nov 14th (cont.)	M75 "Chute" Down @ 37cm	Nov142003_143603	1
	M75 Side @ 36cm	Nov142003_143730	2
	Timing Calibration	Nov142003_144143	3
	Continue Survey of Open Field	Nov142003_144952	41
		Nov142003_153459	68
		Nov142003_164758	11
		Nov142003_165955	62
	Timing Calibration	Nov142003_182838	3
	Continue Survey of Open Field	Nov142003_183356	44
		Nov142003_193029	60
		Nov142003_203645	81
		Nov142003_220233	49
	Timing Calibration	Nov142003_225329	3
Mon, Nov 17th	Warm Up	Nov172003_133515	14
	Ferrite Calibration	Nov172003_134933	4
	4" Al Sphere Calibration	Nov172003_135435	4
	Timing Calibration	Nov172003_135916	2
	Continue Survey of Open Field	Nov172003_140411	62
		Nov172003_151017	61
		Nov172003_161438	67
		Nov172003_172548	42
		Nov172003_181404	26
		Nov172003_184327	65
		Nov172003_195157	20
		Nov172003_204137	3
		Nov172003_210129	60
		Nov172003_220455	37
	Timing Calibration	Nov172003_224817	3
Tues, Nov 18th	Warm Up	Nov182003_172218	17
	Timing Calibration	Nov182003_173959	4

Date	Description	Survey Base Name	Duration (min)
Tues, Nov 18th (cont.)	Continue Survey of Open Field	Nov182003_174907	61
		Nov182003_185255	68
		Nov182003_200050	1
		Nov182003_201016	3
		Nov182003_203400	61
		Nov182003_213808	57
	Timing Calibration	Nov182003_223539	2
Wed, Nov 19th	Warm Up	Nov192003_133215	15
	Ferrite Calibration	Nov192003_135303	4
	4" Al Sphere Calibration	Nov192003_135645	4
	Timing Calibration	Nov192003_140144	2
	Complete Survey of Open Field (Fill-in surveys around trees, wash, fence, etc.)	Nov192003_140730	13
		Nov192003_143539	4
		Nov192003_144107	3
		Nov192003_144710	7
		Nov192003_145917	9
		Nov192003_151011	7
		Nov192003_151849	8
		Nov192003_152715	1
		Nov192003_152930	15
		Nov192003_155222	20
		Nov192003_161656	17
		Nov192003_163806	4
		Nov192003_164328	6
		Nov192003_165814	34
		Nov192003_173521	17
		Nov192003_175504	5
		Nov192003_180541	24
		Nov192003_184518	9
		Nov192003_185835	2
		Nov192003_190040	9
		Nov192003_191111	13
		Nov192003_192729	4
Total Survey Time			**1642**
Total Survey Plus Calibration Time			**1798**

Notice that almost an entire day of surveying was devoted to clean-up surveys around the wash, the trees in the middle of the open field and the fence. In a commercial survey, these areas would have been covered in parallel with the main survey using a hand-held instrument. As this was a demonstration of our vehicular-towed array, we used the array for the entire site.

The equipment returned from the Yuma survey in reasonably good shape. We had to rethink our strain relief for the sensor cables as we lost two cables during the course of the Yuma deployment. These were re-engineered and ready to go shortly after we returned. We delayed the final survey of the Aberdeen site until we had received our scores from Yuma, in case any major changes to our equipment or procedures was required. We returned to Aberdeen in June 2004 to survey the Open Field site. Details of that survey are given in Table 6.

Table 6. Details of the 2004 survey of Aberdeen.

Date	Description	Survey Base Name	Duration (min)
Mon, Jun 7th	Ferrite Calibration	Jun072004_115042	3
	4" Al Sphere Calibration	Jun072004_115542	4
	Timing Calibration	Jun072004_115935	1
	NS Survey of Open Field (interleaved tracks to obtain 0.25-m lane spacing)	Jun072004_120437	58
		Jun072004_130816	51
		Jun072004_140153	57
		Jun072004_151810	66
		Jun072004_163054	56
		Jun072004_172926	22
Tues, Jun 8th	Warm Up	Jun082004_083341	13
	Ferrite Calibration	Jun082004_084743	3
	4" Al Sphere Calibration	Jun082004_085140	4
	Timing Calibration	Jun082004_085624	3
	Continue NS Survey of Open Field (interleaved tracks to obtain 0.25-m lane spacing)	Jun082004_090144	68
		Jun082004_101431	46
		Jun082004_111936	62
		Jun082004_122419	56
		Jun082004_132627	60
		Jun082004_143008	46
		Jun082004_151851	54
	Ferrite Calibration	Jun082004_161509	3
	4" Al Sphere Calibration	Jun082004_161842	3
	Timing Calibration	Jun082004_162154	3

Date	Description	Survey Base Name	Duration (min)
Wed, Jun 9th	Warm Up	Jun092004_074926	21
	Ferrite Calibration	Jun092004_081503	3
	4" Al Sphere Calibration	Jun092004_081827	3
	Timing Calibration	Jun092004_082136	3
Wed, Jun 9th (cont.)	Finish Survey of Open Field (interleaved tracks to obtain 0.25-m lane spacing)	Jun092004_082749	57
		Jun092004_093101	17
		Jun092004_095802	52
		Jun092004_105202	2
		Jun092004_111603	55
		Jun092004_130121	63
		Jun092004_140819	26
		Jun092004_163800	2
		Jun092004_164138	6
		Jun092004_165018	15
	Ferrite Calibration	Jun092004_162824	2
	4" Al Sphere Calibration	Jun092004_163038	3
	Timing Calibration	Jun092004_163330	3
Total Survey Time			**997**
Total Survey Plus Calibration Time			**1041**

3.4.3 Area Characterized

The areas of the two test sites and the MTADS GEM array coverage are detailed in Table 7.

Table 7. MTADS GEM array coverage at the two test sites.

Site	Scenario	Area (ha)	MTADS GEM array coverage (ha)	Fraction
Aberdeen	Calibration Lanes	0.12	0.12	100%
	Blind Grid	0.19	0.19	100%
	Open Field	5.54	4.89	88%
Yuma	Calibration Lanes	0.11	0.11	100%
	Blind Grid	0.17	0.17	100%
	Open Field	6.23	5.92	95%

3.4.4 Operating Parameters of the Technology

The MTADS GEM-3 array is a vehicular-towed array with the operating limitations of a towed system. Support requirements are not extensive. For normal operation, we store the vehicle and sensor trailer in a garage or similar space overnight; this has often been a two-door Conex trailer. We require power at the storage site to charge the vehicle batteries overnight. Other batteries (sensors, GPS base station, radios, etc.) can be charged off-site or in a hotel room but, of course, it is most convenient to charge them at the storage site. For extended surveys we procure an office space for the data preprocessor and to serve as electronics diagnosis and repair area. All of this support was provided at the Standardized Test Sites so no action on our part was required.

The minimum survey crew consists of a vehicle driver, one other field helper (for safety considerations), and a data QC and preprocessing analyst. The surveys described in this report included an additional member of the team who split his time between the field, data preprocessing, and general troubleshooting. As the data preprocessor has time available, target analysis can begin in the field. The bulk of target analysis is carried out later at the analyst's office.

3.4.5 Experimental Design

The processing steps necessary to consolidate the individual raw data files recorded by the data acquisition computer (refer to section 2.1) to a product feasible for analysis are coarsely outlined in the block diagram of Figure 12.

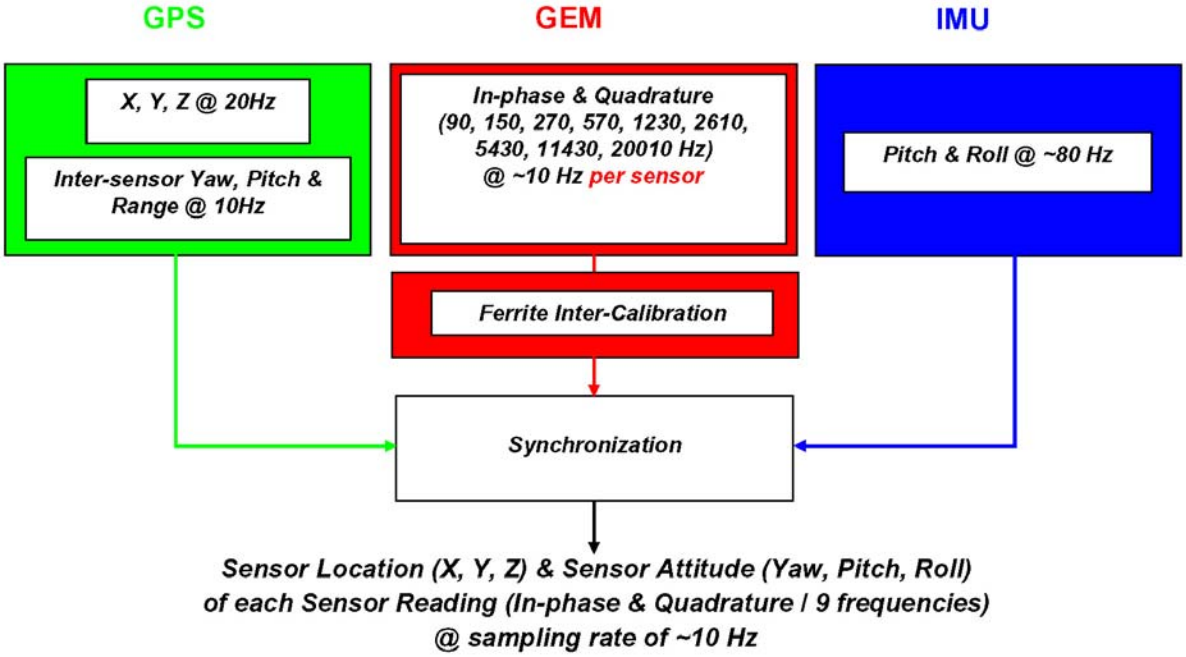

Fig. 12 – A schematic diagram showing the integration of the GEM and navigation data for eventual analysis

Corrections based on recent or same day ferrite measurements are first applied to each of the three GEM data files. In theory, the frequency response to ferrite should be flat and negative for the in-phase, and zero for the quadrature. Amplitude and phase corrections based on this are therefore applied to correct for and equalize the response of all three GEM sensors. In addition, a running median window is subtracted from each of the GEM time series (on a per frequency basis) to mitigate sensor drift – a problem especially at the higher frequencies. A window width of 500 points (~50 s) was chosen for both the APG and YPG Open Field GEM data.

As described in section 2.1, the GPS data files contain NMEA GGK sentences yielding the position of the master antenna (MB1 in Figure 13) at 20 Hz, as well as AVR sentences yielding the vectors to the other two antennae (represented by the red arrows in Figure 13) at 10 Hz. These vectors are defined by their magnitudes (range) and two angles (yaw and tilt), with the yaw angle defined relative to True North. By first redefining the yaw angle to be relative to Grid North, the along-platform and cross-platform vectors are subsequently computed from the two AVR vectors allowing the pitch and roll angles of the platform to be readily determined. Because all sensors are mounted on the same rigid board, the orientation of all the sensors are now completely determined.

To perform the final integration of data, all sensors need to be synchronized. As discussed in sections 2.1 and 2.2.1, this is mostly taken care of in the data acquisition phase by referencing all data to the GPS 1-PPS, with an added sensor latency determined by driving the array back-and-forth over a metal pipe. Since this latter latency tends to change unpredictably over time, it is necessary to perform such a measurement on a continual basis. On implementation of this latency, the location and orientation – a combination of the IMU and GPS pitch and roll (section 2.2.2) – of each GEM sensor at a given time is matched to a given spectrum.

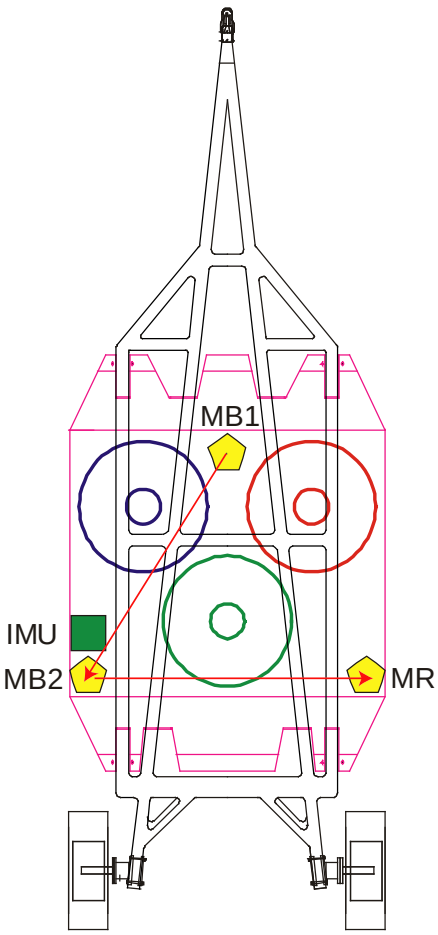

Fig. 13 – The sensor platform

3.4.6 Demobilization

MTADS equipment was removed from the site in the same manner it arrived, roll-back trucks and rental van in the case of Aberdeen and rented 53' truck with government-contract transportation for the survey in Yuma.

4. Performance Assessment

The primary performance metrics of this demonstration are the detection and discrimination performance of the MTADS GEM array at the two Standardized UXO Technology Demonstration Test Sites. These results are provided by the site managers after detection and discrimination picks are submitted [8-11].

4.1 Aberdeen Proving Ground Blind Grid

We will begin our discussion of the demonstration results by concentrating on the APG Blind Grid results. This was the first scoring report we received and it is the data set that we have devoted the most analysis time to understanding. A number of the discrimination strategies we used on these data were modified based on the scoring report and the updated strategies applied on the other scenarios and sites.

4.1.1 Overview of the Survey Data

Anomaly images of the survey data collected at APG at two representative frequencies are shown in Figures 14 through 17. The data at 150 Hz are representative of the lower frequencies with relatively high noise levels in the in-phase response and much less noise in the quadrature response. The data collected at 11,430 Hz is representative of the higher frequencies with less noise in the in-phase response and less signal in the quadrature response for a number of the objects in the Blind Grid.

There are several points to be made about these data. The 150 Hz in-phase data, and to a lesser extent the 11,430 Hz in-phase data, show the effects of sensor drift. Drift is always an issue with these sensors and we routinely correct for this drift by applying a running median window of 500 points as described in Section 3.4.5. The density of targets in the Blind Grids at both of the Standardized Test Sites does not allow this drift correction method. For the blind grids, we used data from the clear areas above and below the Blind Grid to level the measured data. This obviously corrects well for longer-term drift but less well for shorter term drift. This drift is not such an issue in the quadrature channels as can be seen in Figures 15 and 17.

At the time of our survey, there was a large puddle of standing water on the NE corner of the site. The Open Field part of the site was so wet that we were unable to survey it in the first deployment. The effects of the standing water can be seen in the in-phase response at both frequencies shown. It is so strong that it obscures the response of the shallow shotputs that are buried along the edges of the Blind Grid

For comparison, the 11,430 Hz in-phase response measured at the YPG Blind Grid is shown in Figure 18 using the same intensity scale. The geologic noise at the Yuma site is both higher amplitude and on a shorter length scale than that at Aberdeen. This has obvious implications for the ability of this, or any multi-frequency EMI system to classify targets.

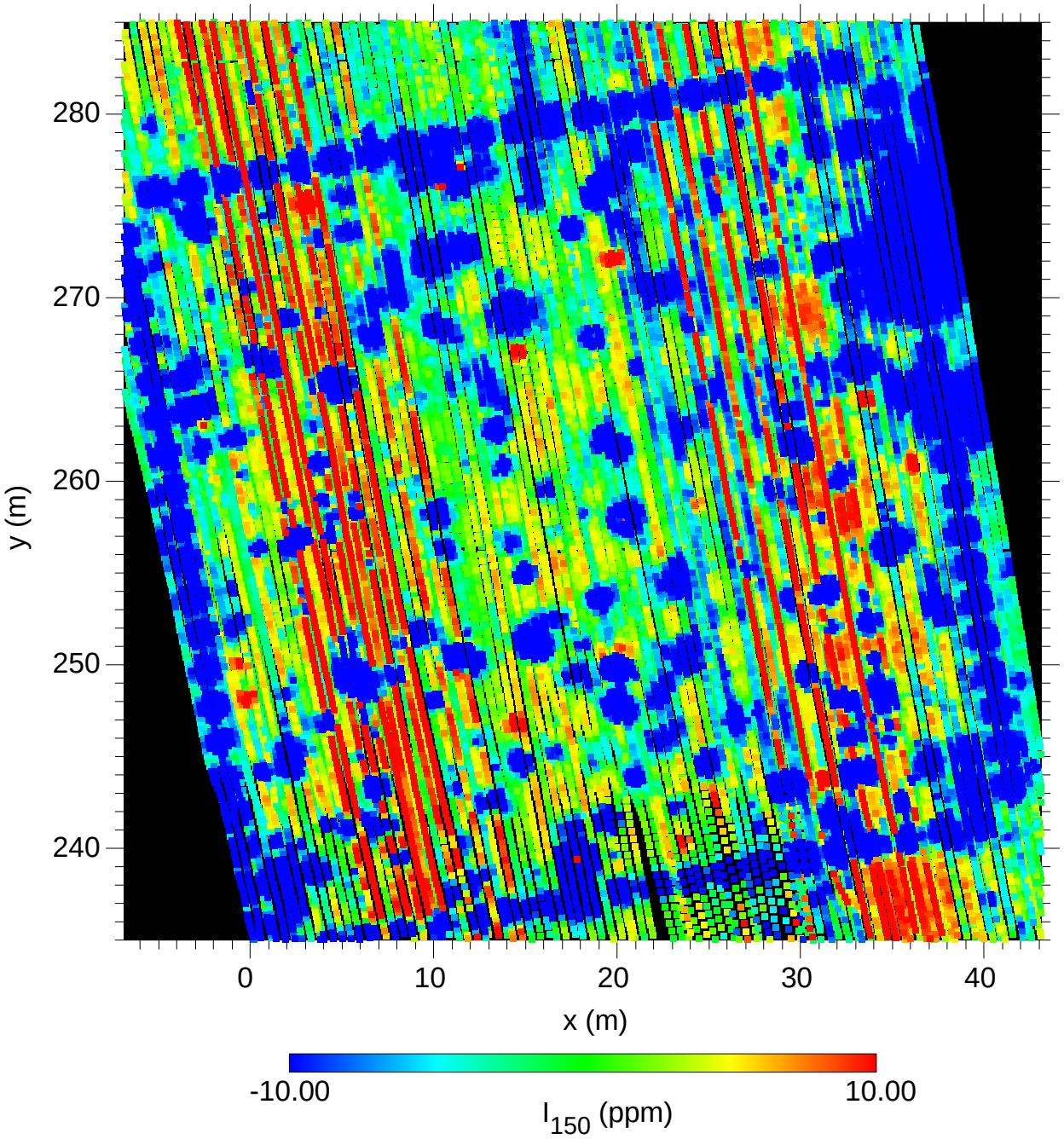

Figure 14 – Anomaly image of the 150Hz in-phase response at the APG Blind Grid

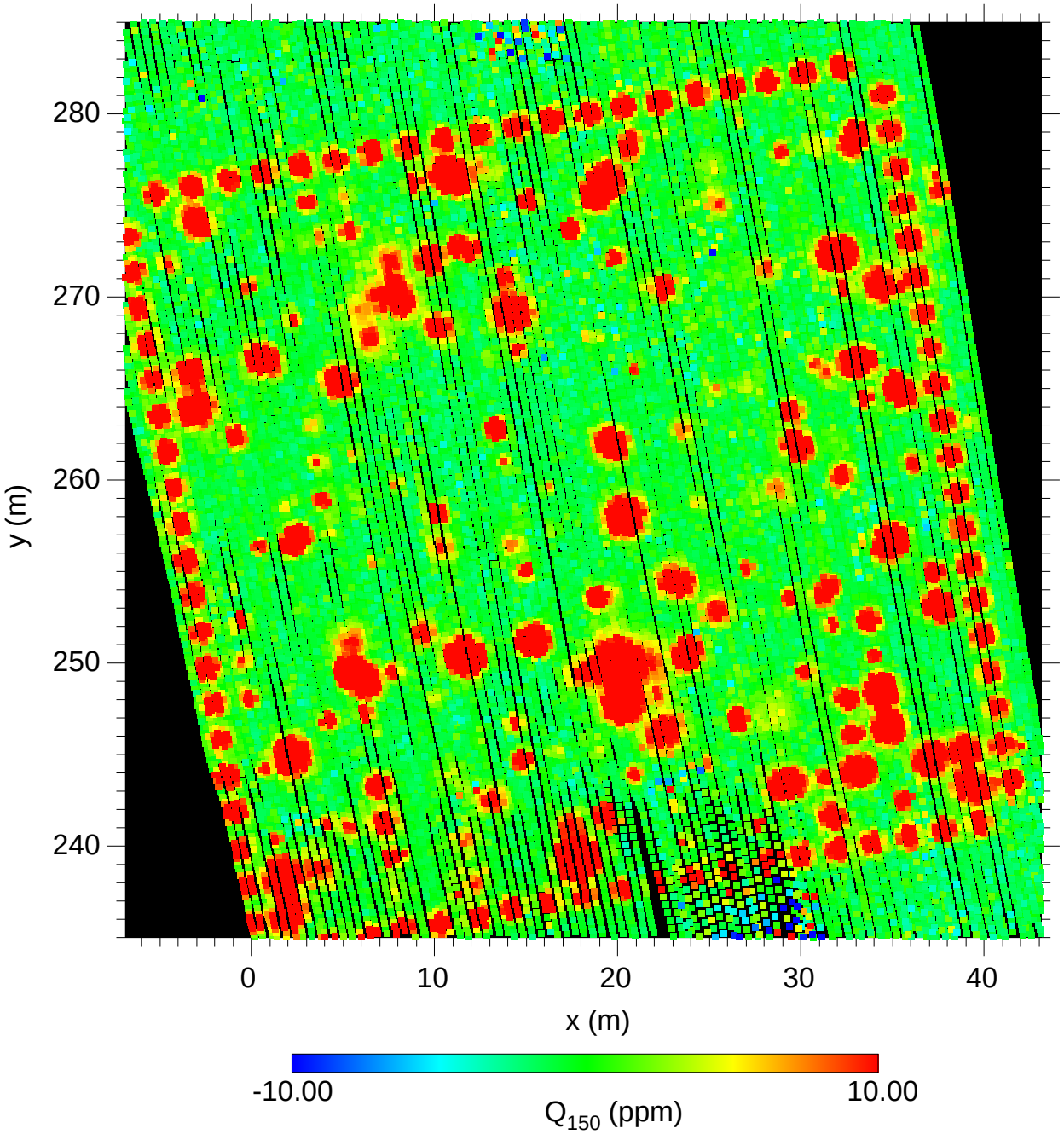

Figure 15 – Anomaly image of the 150 Hz quadrature response at the APG Blind Grid

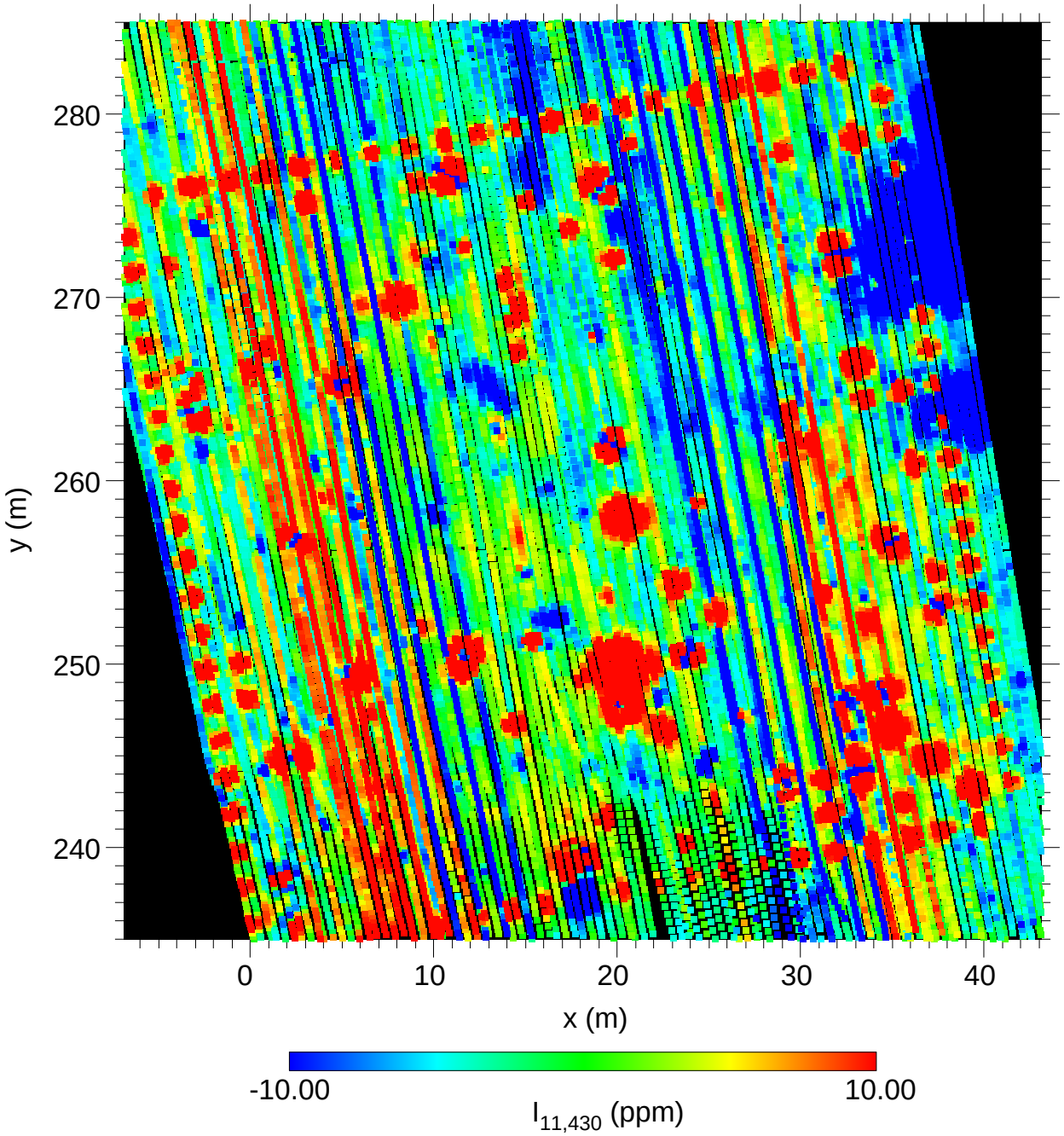

Figure 16 – Anomaly image of the 11,430 Hz in-phase response at the APG Blind Grid

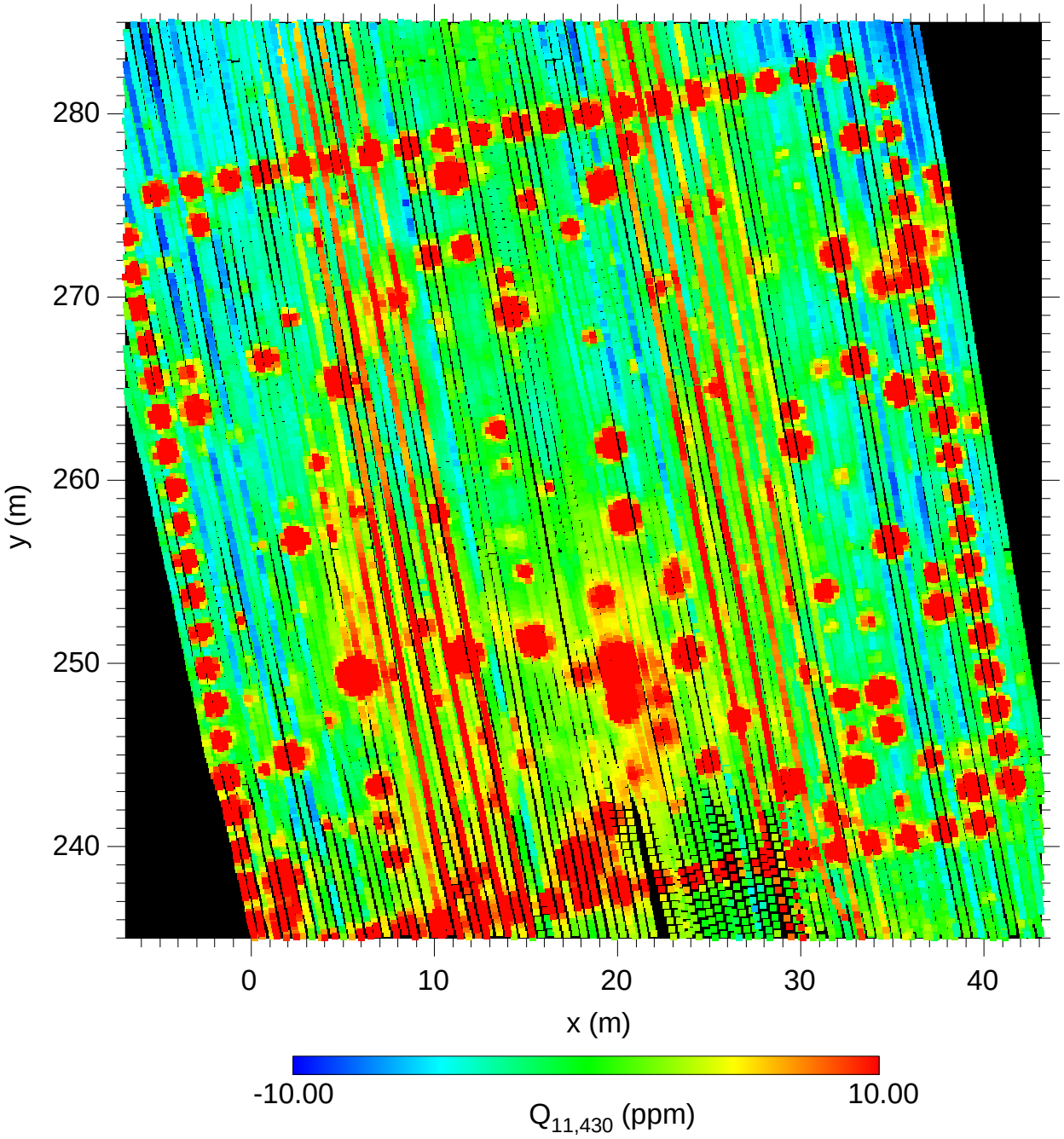

Figure 17 – Anomaly image of the 11,430 Hz quadrature response at the APG Blind Grid

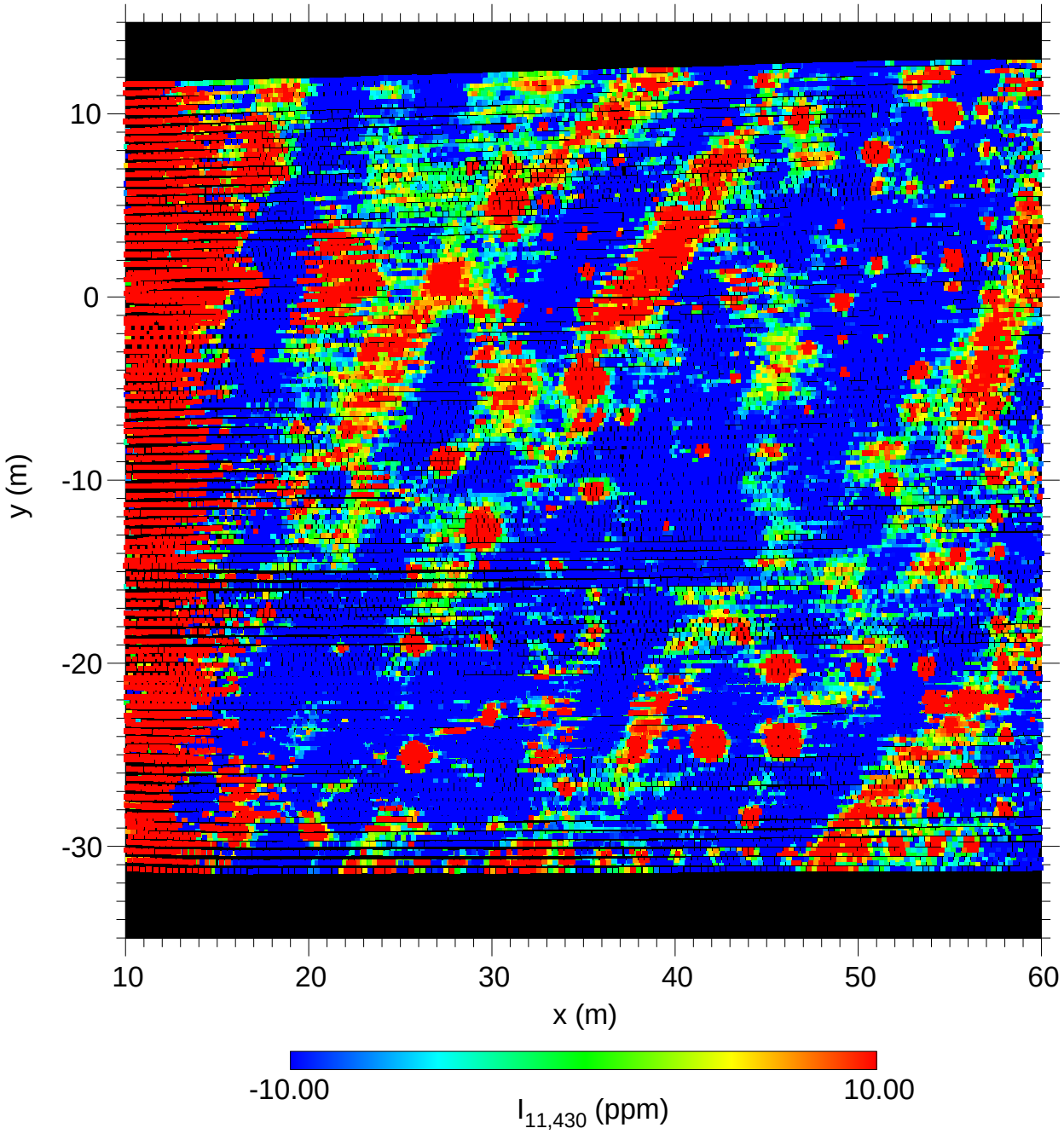

Figure 18 – Anomaly image of the 11,430 in-phase response at the YPG Blind Grid

4.1.2 Response Stage

The first stage of scoring at the Test Sites is the Response Stage. For this, we use the Q_{avg} quantity; the average of the quadrature response for the middle five frequencies. We choose this metric because of the lower noise in the quadrature response and the good signal in the mid frequencies for the objects of interest. A Q_{avg} plot for the APG Blind Grid is shown in Figure 19.

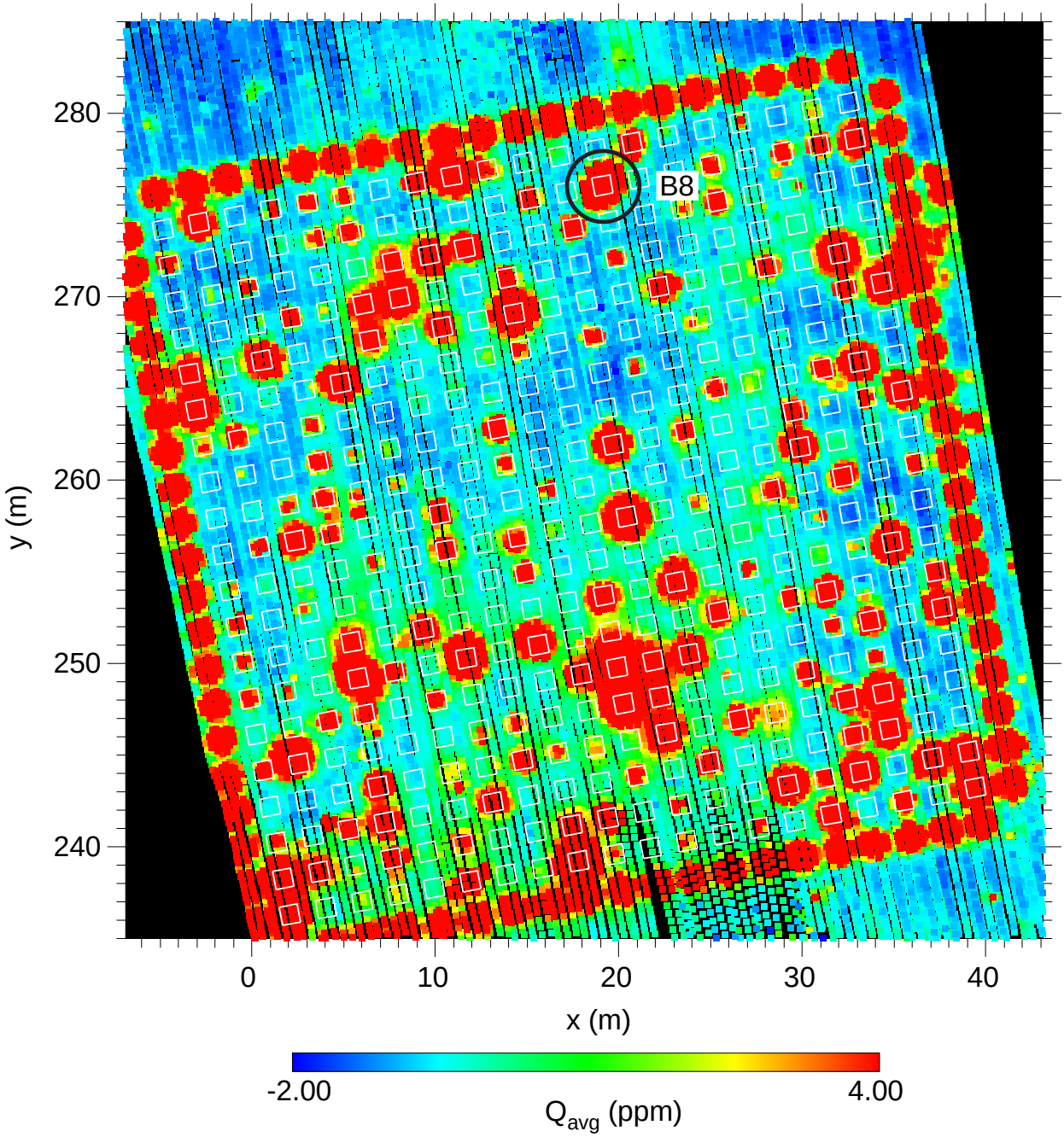

Figure 19 – Q_{avg} anomaly image map of the APG Blind Grid

The 400 cells in the Blind Grid are marked with white squares in Figure 19. A summary of the GEM array detection performance is given in Table 8.

Table 8. Summary of Detection Performance at the APG Blind Grid.

Cell Contents	Number of Cells	Number Correct	Number Incorrect
Single Ordnance Item	84	73	11
Ordnance Item with Clutter	7	7	0
Single Clutter Item	95	91	4
Two Clutter Items	8	8	0
"Empty"	206	174	32
Total	**400**		

The 32 cells reported as "Empty" but for which we made a declaration require some discussion. Only 12 of these false positives showed signal in the GEM array survey only. Seven of these cells had a detection by the GEM array, the EM61 HH, and the magnetometer array. Ten had a detection by the GEM array and the EM61 HH and 3 had a detection by the GEM array and the magnetometer array. An example of this is cell B8 which is highlighted in Figure 19. It is difficult to understand the observed signal unless there is some inadvertent metal in this cell.

An indication of the depth performance of the system is shown in Figure 20. The detected items are shown as black triangles and the missed items are shown as red crosses. The reference line corresponds to a depth of 11x the item diameter. As can be seen, the GEM array is capable of detecting targets down to and below 11 times their diameter at this site.

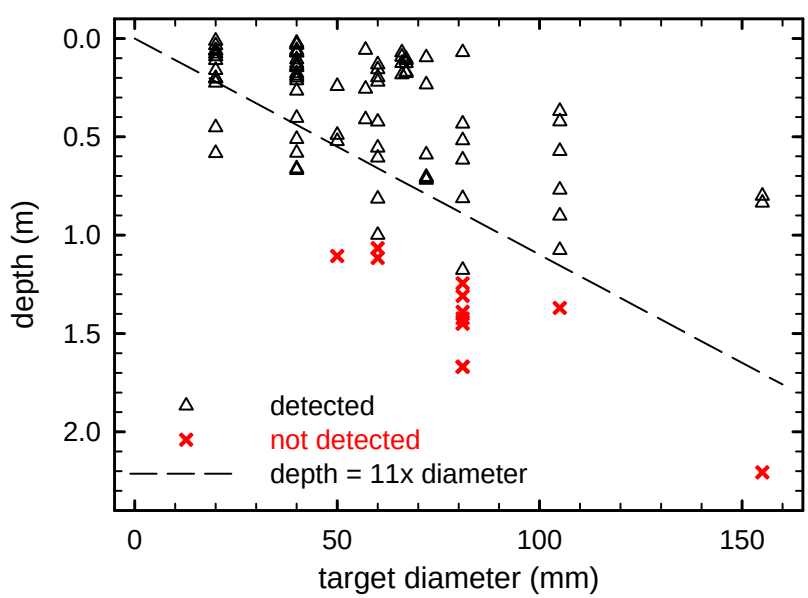

Figure 20 – Detection performance as a function of depth at the APG Blind Grid

The response stage data is shown plotted in Figures 21 and 22 in the manner of the Standardized Test Site scoring reports. Figure 21 shows cumulative ordnance count vs cumulative clutter count. Since the targets are ordered by signal amplitude at the response stage it is no surprise that this plot is essentially along the diagonal. A better measure of system capability is shown in Figure 22 which plots cumulative occupied cells vs adjusted cumulative blank cells. Cells such as B8 which obviously contain buried metal were excluded from the blank count.

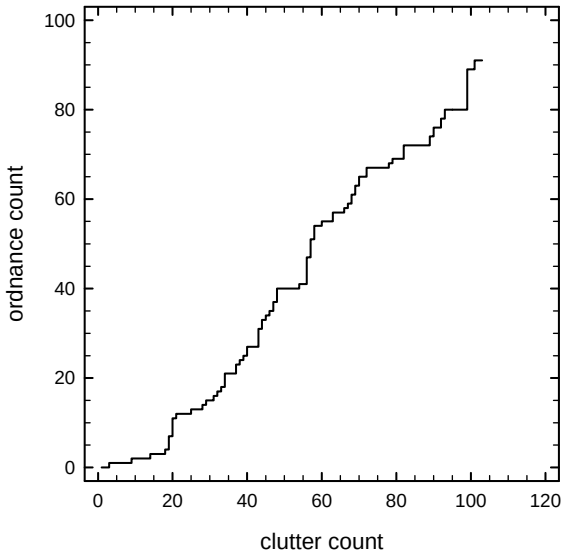

Figure 21 – Response stage results showing cumulative ordnance count vs cumulative clutter count

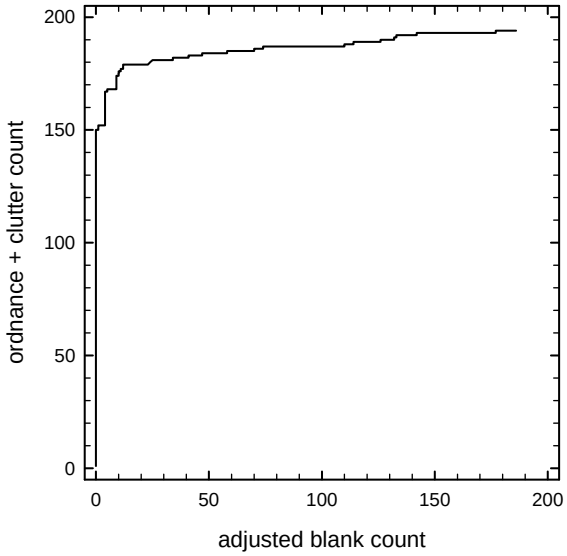

Figure 22 – Response stage performance showing cumulative occupied cell count plotted vs adjusted cumulative blank cell count

4.1.3 Discrimination Stage

An obvious first choice for discrimination methodology at the Standardized Test Sites is library matching since the target set is limited and examples are available to compile a good library of responses. Accordingly, we obtained a set of the targets from the Aberdeen Test Center and measured their response in three orientations (nose up, nose down, and horizontal). Figure 23 shows a 60-mm mortar under test and Figure 24 shows the resulting measured responses for the three orientations.

Figure 23 – Library response data being acquired for a 60-mm mortar

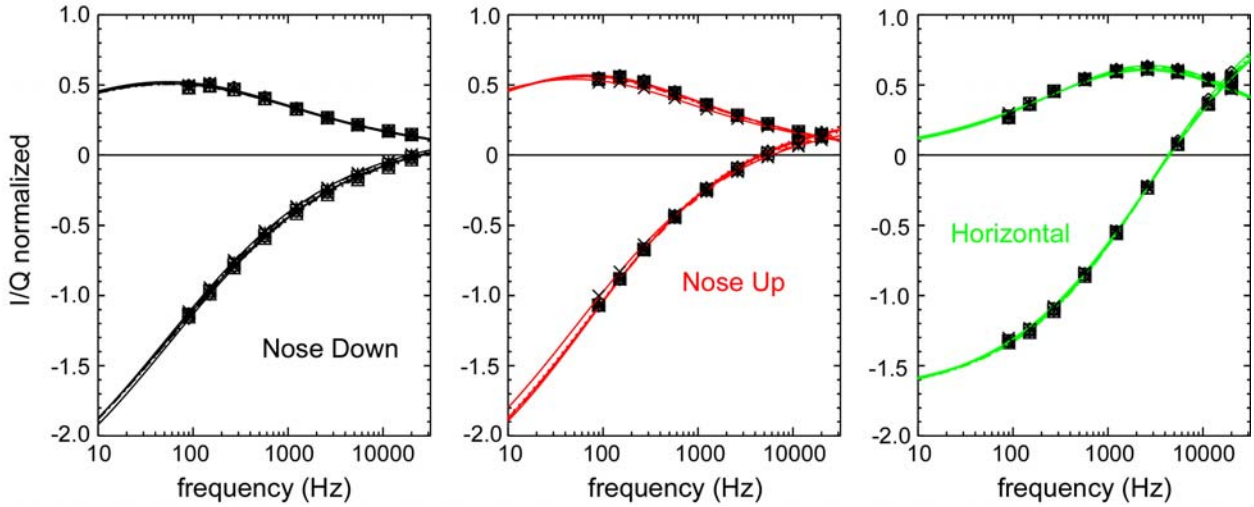

Figure 24 – Measured response curves for the mortar pictured in Figure 23

Having obtained a (nearly) complete set of library response functions, the prescription for discrimination is to compare the measured response of a target to each library set in turn and determine which library item results in the closest match and determine if the match is close enough to declare the target as a member of the library. For this, we compute the χ^2 for the best match. Of course, if the signal level for the various targets differs by a large amount (some targets quite shallow and some very deep) the computed χ^2 can be strongly affected by the signal amplitude. To test this possibility, we computed the χ^2 for the best match weighting the data by the usual $1/\text{rms}^2$ (where the rms deviation is determined from areas between targets) and by 1/(rms + 0.01 x signal) in an attempt to reduce the influence of depth on the computed χ^2.

Figure 25 shows the results obtained for both weighting methods for the targets in the APG Calibration Lanes, the APG Blind Grid, and test data we took in the Blossom Point Test Pit. In all three cases, the computed χ^2 is plotted versus item number with the ordnance items denoted by diamonds and the clutter items denoted by x's. The solid line on the plots is the log-mean of all ordnance in the APG Calibration Lanes with the mean plus one and two standard deviations denoted by the dashed lines.

The χ^2 calculated with the signal-based weighting was used for our declarations at the APG Blind Grid. Based on the results from the Calibration Lanes (which was all we had available at the time), we established a χ^2 threshold of 0.01 for the ordnance/clutter decisions. This is a little less than three standard deviations above the ordnance mean. The reported Discrimination Response Factor was just the inverse of the χ^2.

As can be seen from the right-hand side of Figure 25, a threshold of 0.01 results in one false negative in the Calibration Lanes and correctly excludes a good fraction of the clutter. The results were not as positive at the Blind Grid. There were a number of false negatives and more than half of the clutter items were declared as ordnance. Of greater concern are the results of applying this analysis to data collected later at the Blossom Point Test Pit. These data should be substantially cleaner than the other two sets as the area around the test pit is relatively smooth and there should be a minimum of angular and vertical variation in the sensor positions.

Figure 26 shows the same six cases with χ^2 plotted against S_{max} rather than item number. Here, S_{max} is defined as

$$S_{max} = \sqrt{I_k^2 + Q_k^2}$$

where k is the frequency with the largest S. The results are most striking for the APG Calibration Lanes where the clutter is relatively shallow compared to the ordnance and thus results in larger amplitude signals. Our attempts to mitigate this by adjusting the weighting factor (right-hand side of the plots) was not successful. In short, this method can be thought of as a computationally-intensive depth detector. It is clear in retrospect that the performance of this algorithm at the APG Calibration Lanes is due to the fact that the clutter has stronger signals.

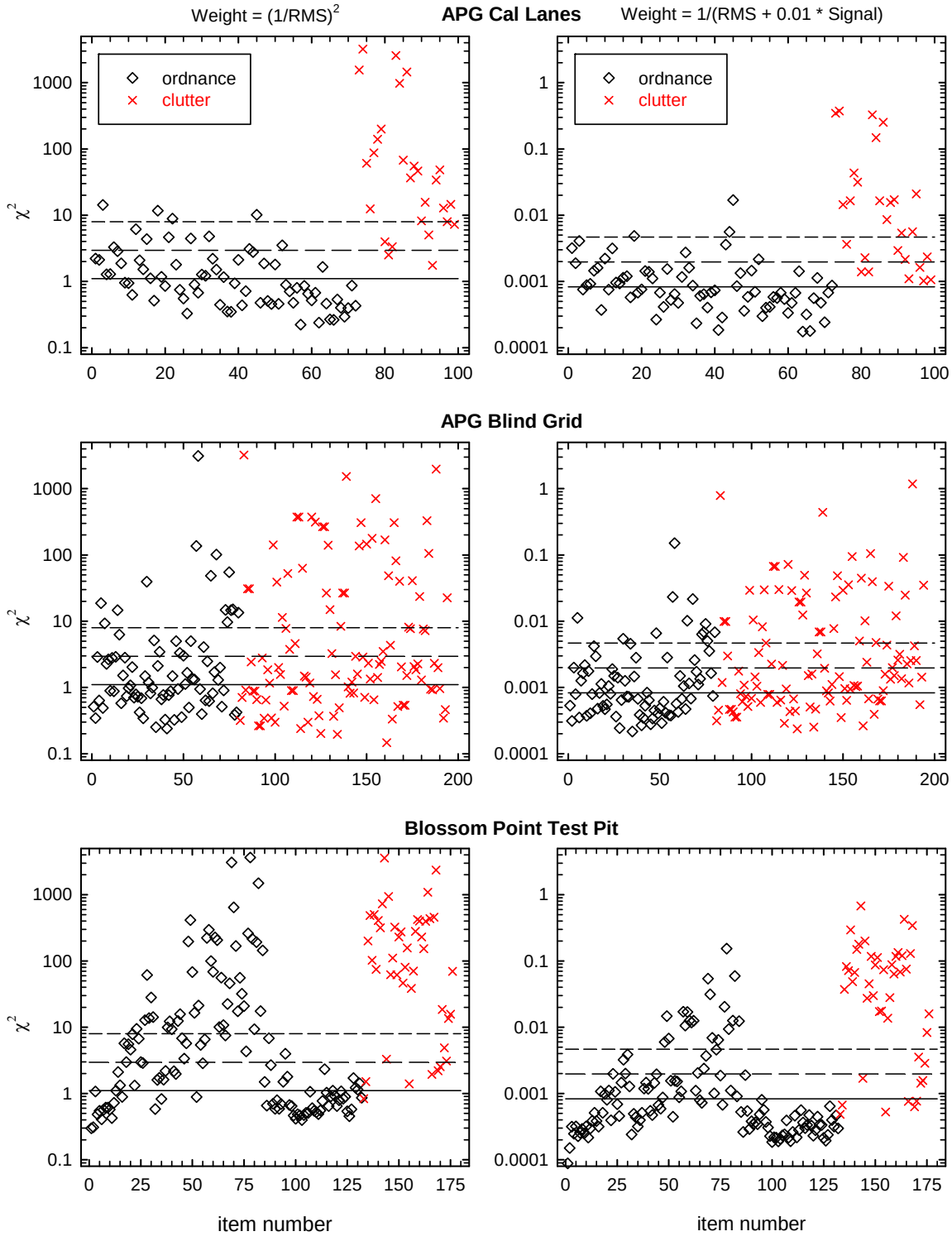

Figure 25 - Plot of χ^2 for best match of target responses to library data for two weighting schemes. See the text for a discussion of the plots.

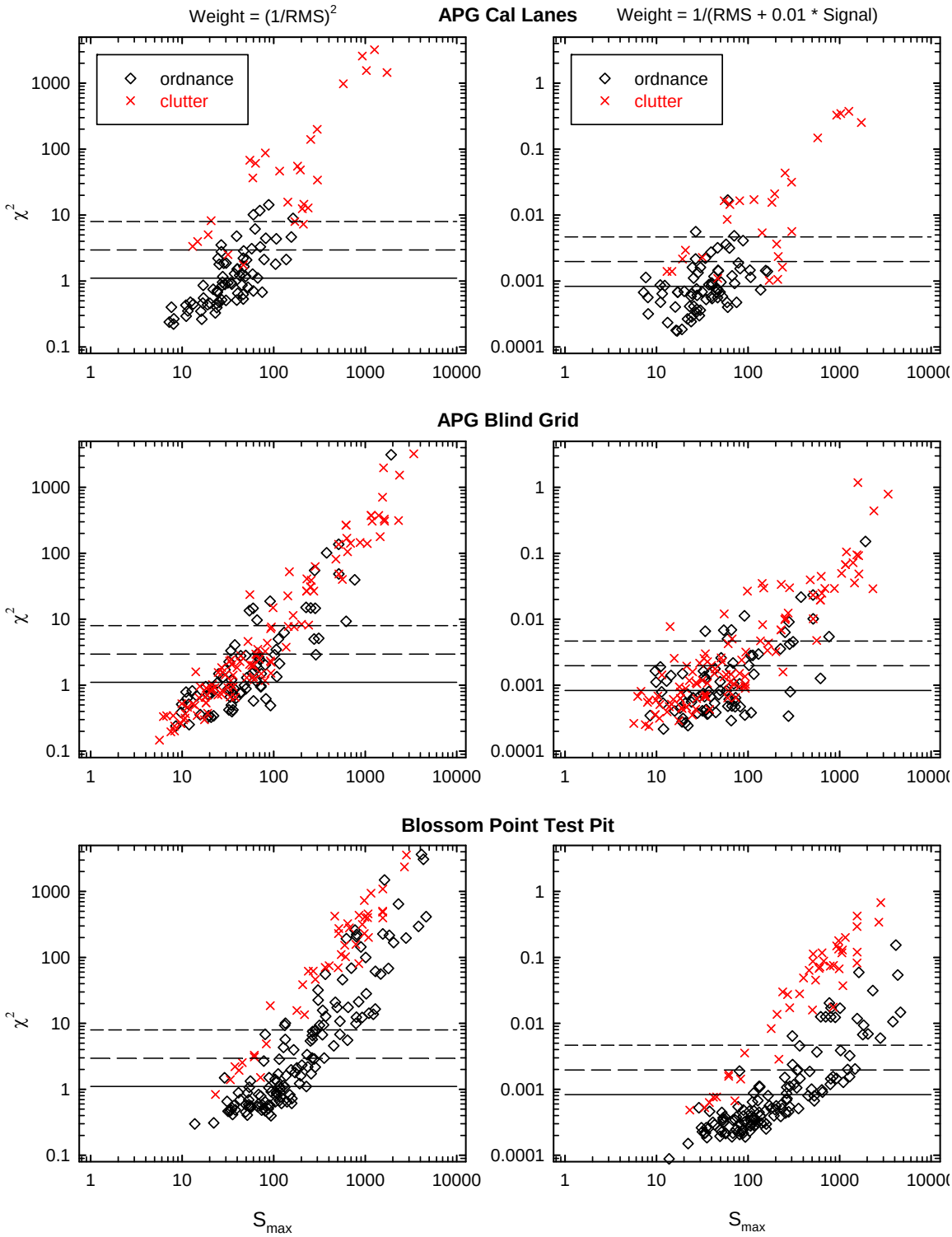

Figure 26 – The data from Figure 25 plotted as a function of maximum signal amplitude rather than item number. See the text for a definition of the terms.

We assume that this strong variation of χ^2 with signal amplitude arises from the bouncing motion of the sensor array as it traverses the rough field. Over a high-signal target, small variations in z result in relative large variations in signal as compared to over a deep, low-signal target. In this case, we can model the bouncing noise by (K * signal) and the correct weighting would be $1/(rms^2 + (K * signal)^2)$. The data from the pit at Blossom Point are shown plotted in Figure 27 as a function of this parameter K. At values of K around 0.3, the scaling of χ^2 with signal amplitude seems to flatten out.

In Figure 28, we have recalculated the best match χ^2 for the three sites using weighting of $1/(rms^2 + (0.3 * signal)^2)$. In this plot, the log-mean and standard deviations for the two APG sites are derived from the Calibration Lanes ordnance while those for the Blossom Point Pit data are derived from the Blossom Point ordnance measurements. Now, as expected, the discrimination algorithm performs best on the higher-quality Blossom Point Pit data.

Scaling the weights by the signal improves the performance of the discrimination but is not very practical as the scaling coefficient is determined after the fact. For the later YPG and APG Open Field demonstrations we employed another method to mitigate the effects of bouncing noise. Each target was fit using a full, unconstrained 3-β model as well as the library model. The ratio of the χ^2 for these two methods, which eliminates the dependence on signal amplitude, should approach 1 if the item is in the library. This method is shown applied to the familiar, original three sites in Figure 29.

The ROC curves for the application of these three discrimination methods to the APG Blind Grid are shown in Figures 30 - 32. The standard χ^2 weighting (Figure 30) and the modified weighting with "bounce noise" added (Figure 31) result in curves that vary little from the chance diagonal. This should not be surprising in the light of Figures 26 and 28. There are fewer items in Figure 31 than in Figure 30. The original submission to APG required that a discrimination score be included for all cells, even those below our detection threshold. We arbitrarily assigned these cells a low discrimination score. The χ^2 with "bouncing noise" analysis was only applied to cells in which we declared a detection.

The χ^2 ratio method (Figure 32) does show some promise. Notice, however, that the curve in Figure 32 includes even fewer ordnance and clutter items than in Figure 31. The χ^2 ratio method requires two different inversions to converge to sensible results in order to calculate the ratio. As the signal-to-noise ratio decreases, this becomes and increasingly difficult hurdle. Library methods such as this can work well when the expected targets are well defined but, of course, can provide inappropriate results when a munition not in the library is encountered.

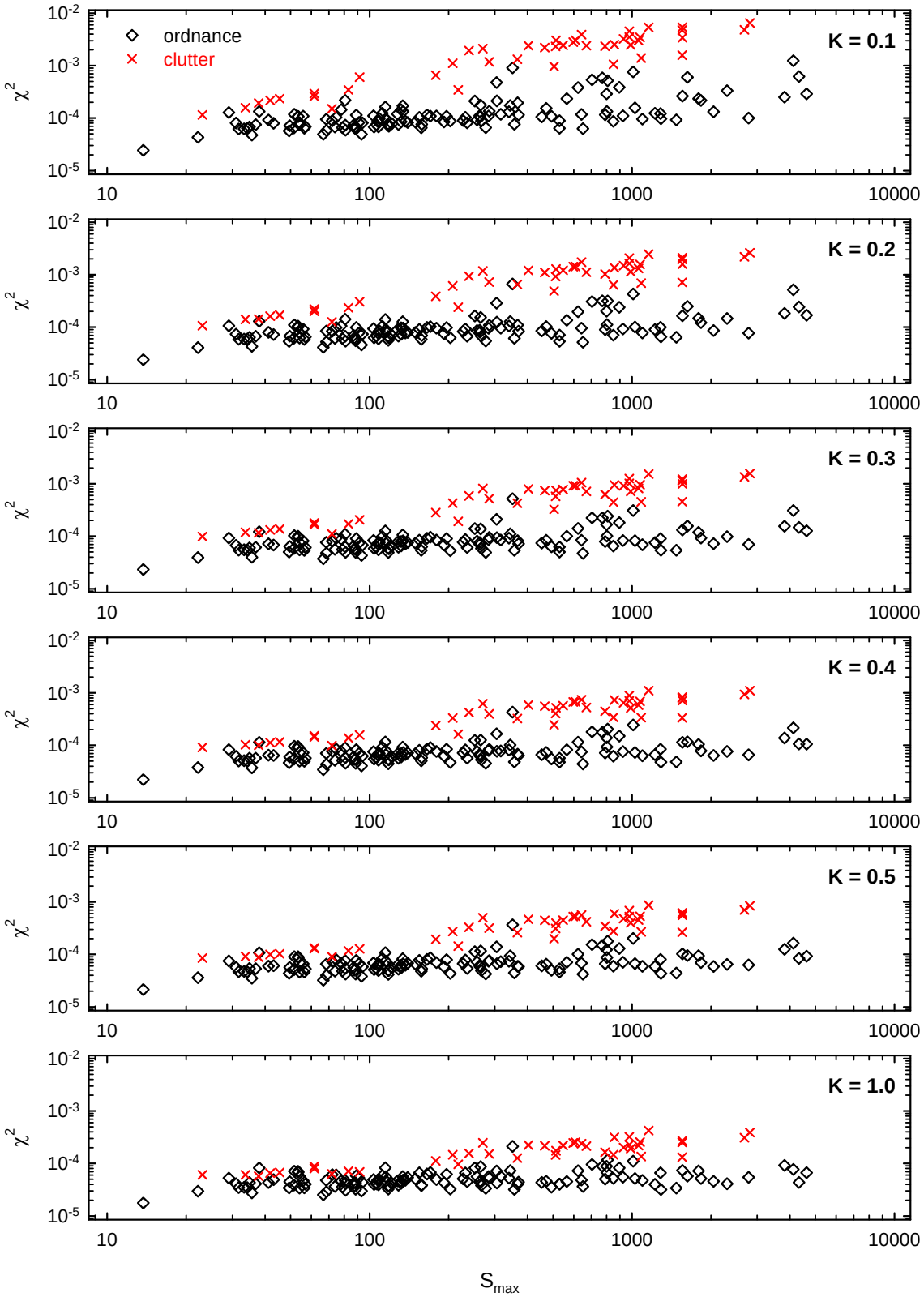

Figure 27 – χ^2 for the Blossom Point Pit data as a function of the amount of "bouncing noise" included in the weighting factor. See the text for a definition of the terms.

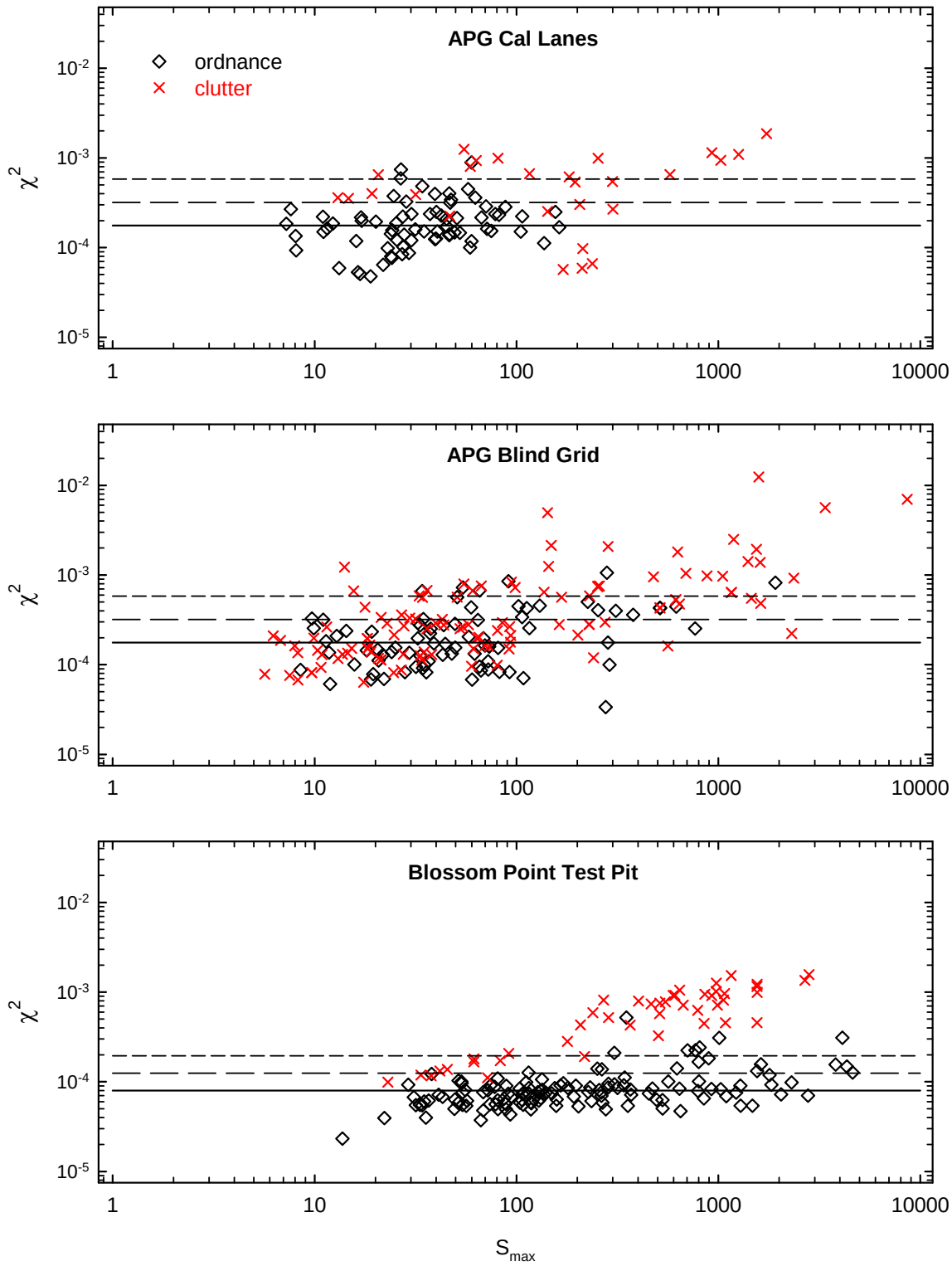

Figure 28 – Recalculation of the best fit χ^2 for the three sites using a value of 0.3 for the "bouncing noise" coefficient as discussed in the text

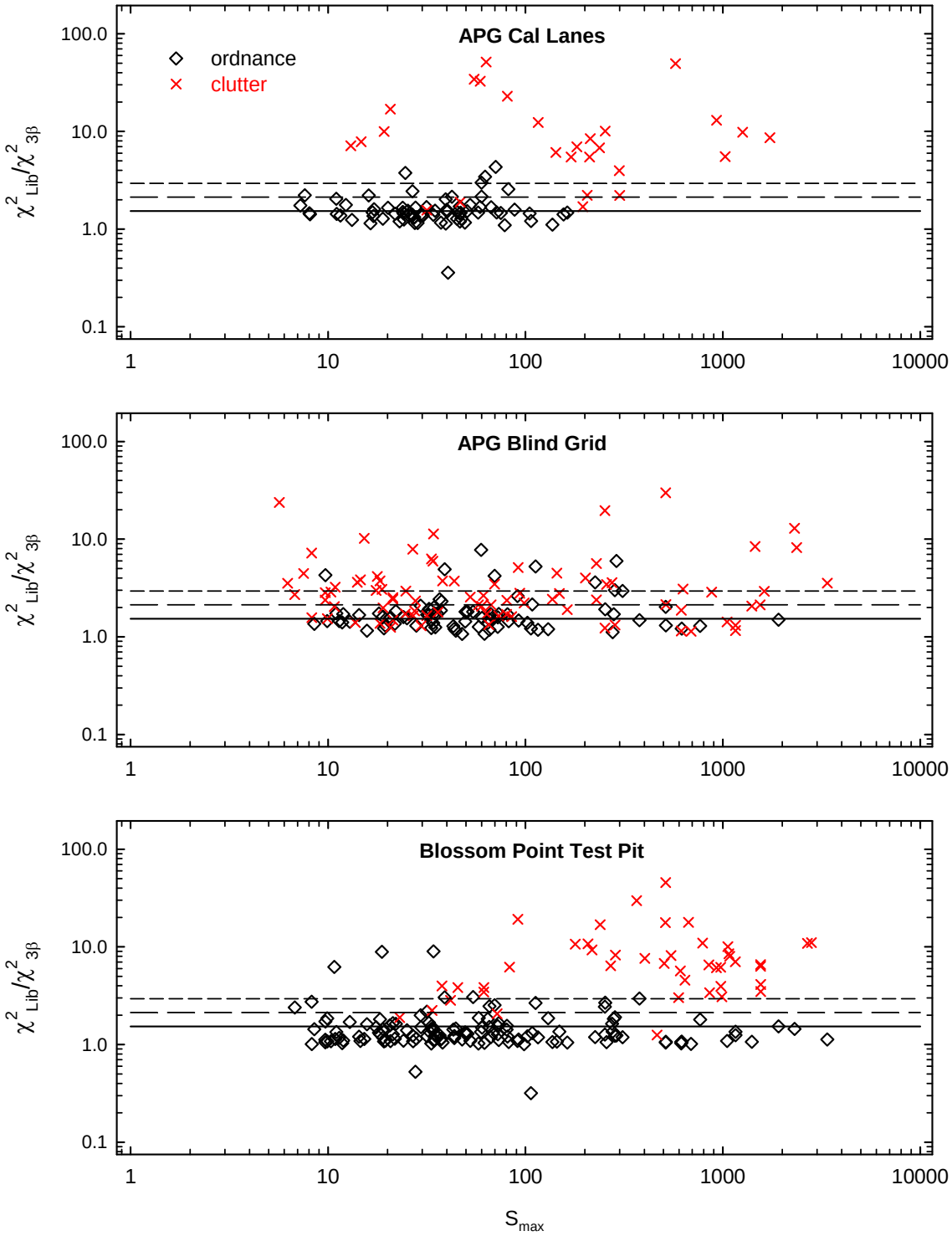

Figure 29 – χ^2 ratio method applied to the data from the three sites that have been discussed in this section

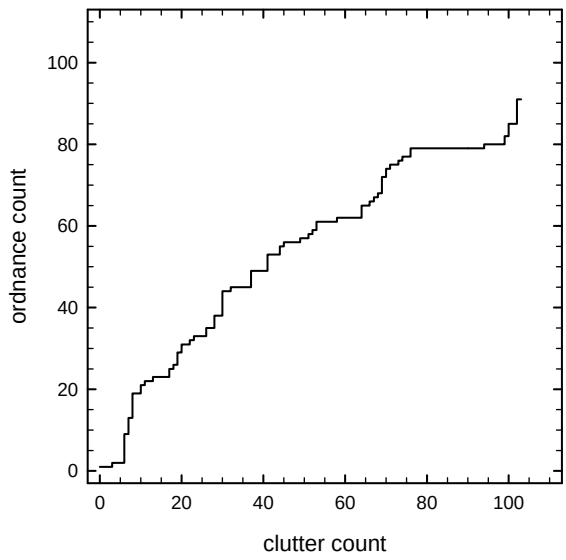

Figure 30 – ROC curve for the χ^2 weighting applied to the APG Blind Grid as shown in the left-hand side of Figures 25 and 26

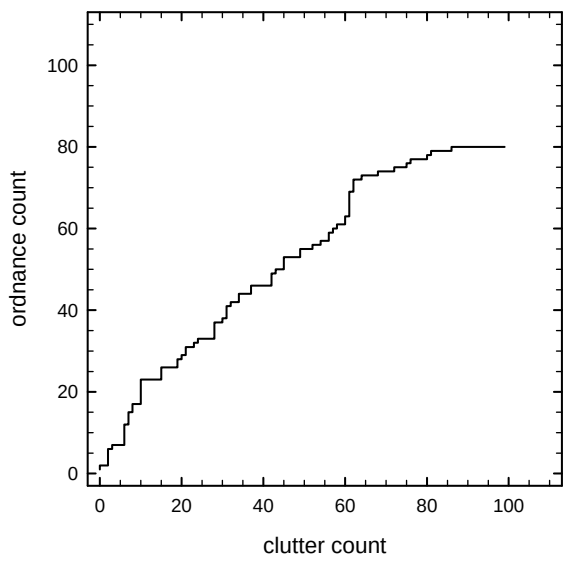

Figure 31 – ROC curve for the case of χ^2 weighting with an estimate of "bouncing noise" included (Figure 28) applied to the APB Blind Grid

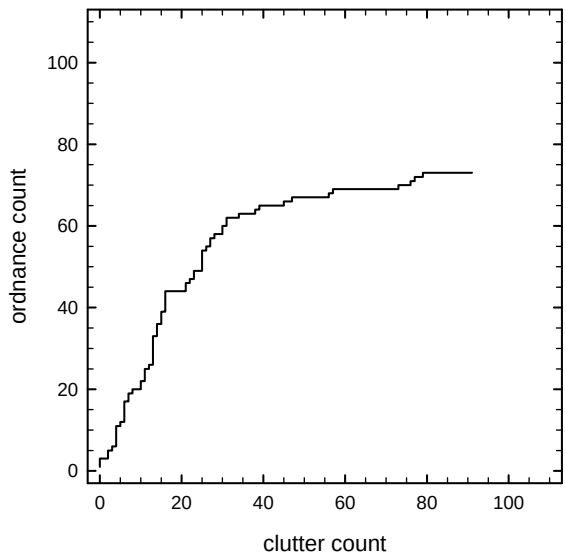

Figure 32 – ROC curve for the χ^2 ratio method (Figure 29) applied to the APG Blind Grid

4.2 Aberdeen Proving Ground Open Field

Selected results from our surveys at the Open Field areas at the two Standardized Test Sites have been provided to us by analysts at the Institute for Defense Analyses. These results are summarized graphically in the following sections.

4.2.1 Response Stage

Response stage data from the Open Field scenario at the APG Standardized Test Site is shown in Figure 33 as a plot of probability of detection vs. normalized background alarm rate. There are analysis models on the plot. The first, the red line in Figure 33, corresponds to considering only those targets that were covered by the survey and are not within 2 m of another target. The analysis corresponding to the blue line retains those limitations and also excludes those targets deeper than 11x their diameter. We showed in Figure 20 that the GEM array demonstrated is able to detect small and medium targets below this relative depth but our detection efficiency falls off at depths below 11x the item diameter. Response stage results broken out by item type are shown in Figure 34. In this figure, the depth of 100% detection is denoted by the blue bar and the depth of maximum detection is shown as the horizontal line. For a number of the items, 105-mm HEAT for example, these two depths are the same. For the majority of the items, the maximum depth of detection is below the depth of 100% detection.

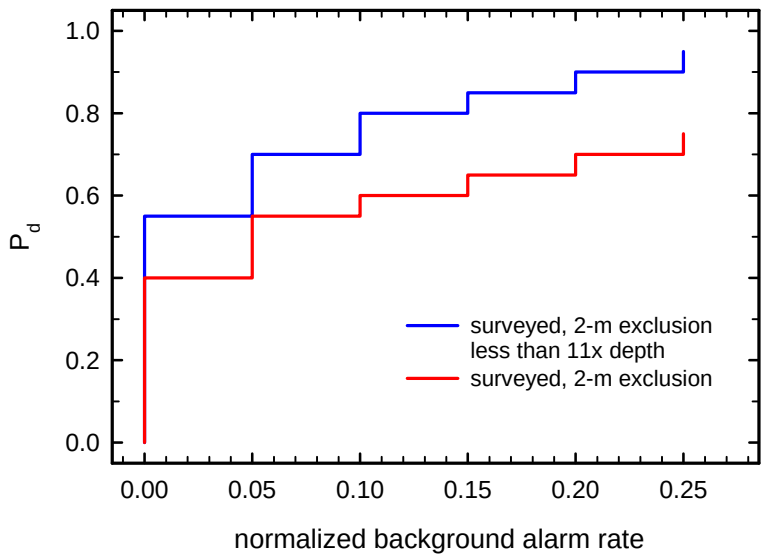

Figure 33 – Detection performance at the APG Open Field Scenario. The red line is derived considering only targets that were covered in the survey and are not within 2 m of another target. The blue line retains those criteria and also excludes targets deeper than 11x their diameter.

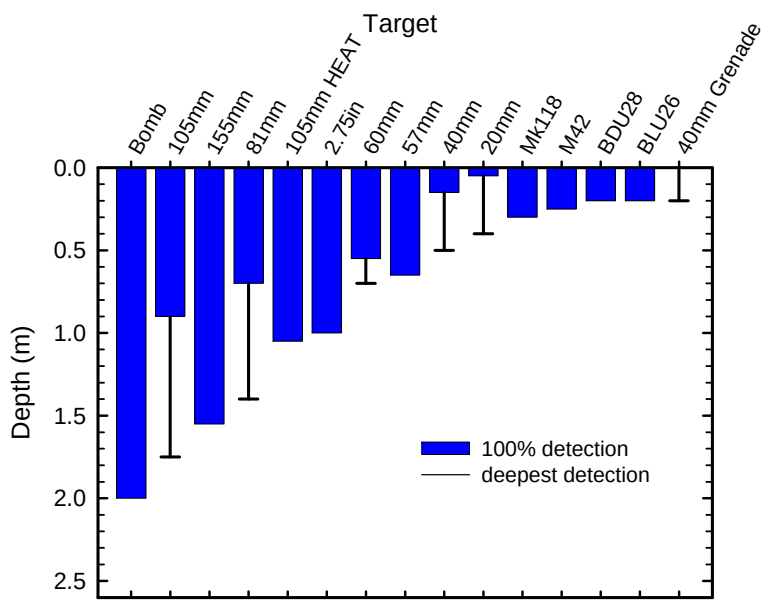

Figure 34 – Response stage results for the APG Open Field scenario broken out by target type

4.2.2 Discrimination Stage

Discrimination stage performance at the APG Open Field using the same two analysis models is shown in Figure 35. As above, the exclusion of items at depths below 11x their diameter (presumably lower S/N anomalies) improves the discrimination performance obtained.

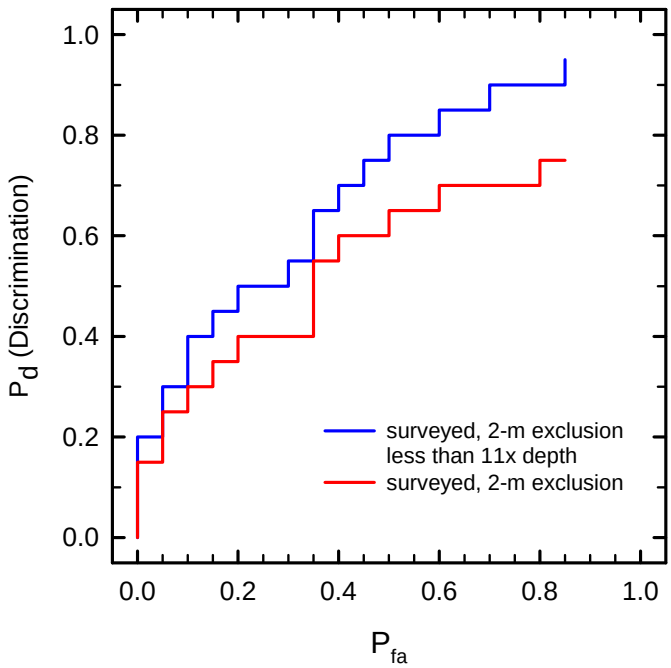

Figure 35 – Discrimination performance at the APG Open Field Scenario. The red line is derived considering only targets that were covered in the survey and are not within 2 m of another target. The blue line retains those criteria and also excludes targets deeper than 11x their diameter.

4.3 Yuma Proving Ground Open Field

4.3.1 Response Stage

Response stage results for the YPG Open Field scenario are shown in Figures 36 and 37. As for APG, they are analyzed by excluding first items that were not covered by the survey or are within 2-m of another item then retaining those exclusions and further excluding items deeper than 11x their diameter. Notice that the background alarm rates in Figure 36 are more than a factor of two smaller than the corresponding results from Aberdeen. Although the Yuma site is more geologically active than Aberdeen, it is smoother so there were fewer false alarms due to platform bouncing over deep ruts. Detection depths at Yuma are, in general, in line with those obtained at Aberdeen. Note however, that a shallow bomb was apparently missed resulting in an unusual plot for that target type.

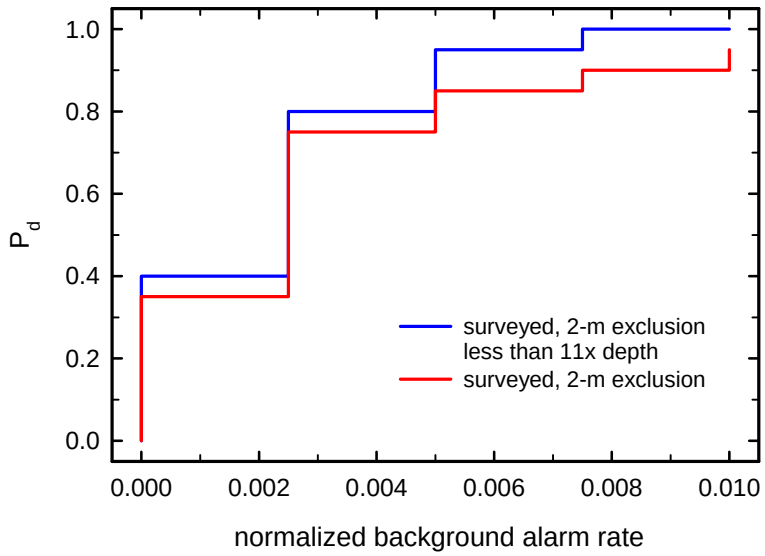

Figure 36 – Detection performance at the YPG Open Field Scenario. The red line is derived considering only targets that were covered in the survey and are not within 2 m of another target. The blue line retains those criteria and also excludes targets deeper than 11x their diameter.

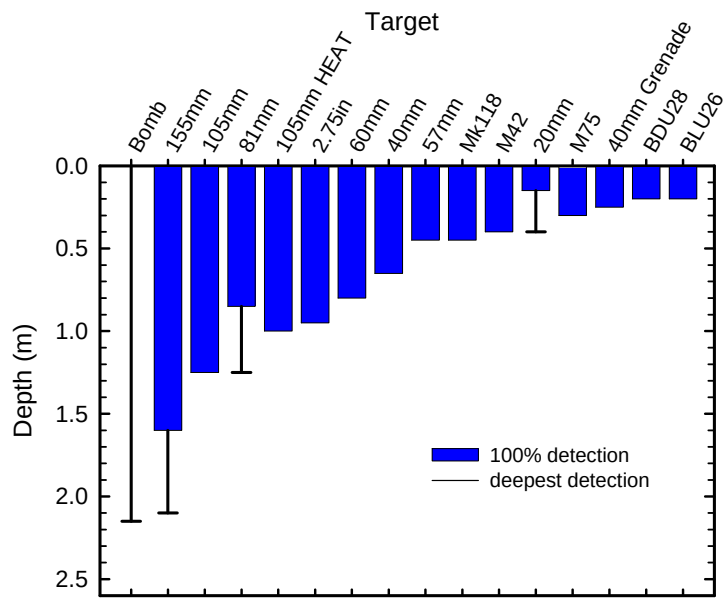

Figure 37 – Response stage results for the YPG Open Field scenario broken out by target type

4.3.2 Discrimination Stage

Discrimination Stage results from the YPG Open Field are shown in Figure 38. As before, exclusion of items that are deeper than 11x their diameter improves performance which is better, on the whole, than that observed at Aberdeen. As with the response stage, this is likely due to the lower platform motion noise observed at the Yuma site.

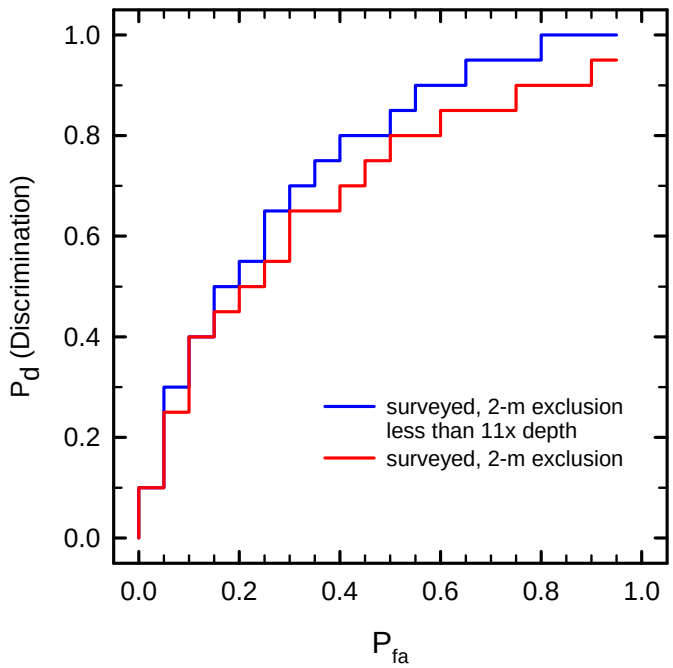

Figure 35 – Discrimination performance at the YPG Open Field Scenario. The red line is derived considering only targets that were covered in the survey and are not within 2 m of another target. The blue line retains those criteria and also excludes targets deeper than 11x their diameter.

5. Cost Assessment

5.1 Cost Reporting

The costs for a 50-acre survey using the MTADS GEM-3 array in discrimination mode are detailed in Table 9. These surveys require interleaved sampling as was practiced in the two Standardized Test Site surveys. The coverage using this collection methodology is ~5 acres/7.5 hour survey day. Thus, a 50-acre survey requires two weeks in the filed. All costs are approximate 2006 costs to make the numbers more meaningful.

Table 9. Summary of Costs for a 50-acre GEM array survey.

Cost Category	Sub-Category	Cost	Sub-Total
Mobilization Costs			**$63,500**
	Preliminary Site Visit	$8,000	
	Test Plan Preparation	$10,000	
	Equipment Prep and Packing/Unpacking	$12,000	
	Rental Trailer and Transportation	$17,500	
	Analysts Set-up	$10,000	
	RT Travel for 4 Personnel	$6,000	
Logistics (if required)			**$13,600**
	Establish GPS Control Points	$4,000	
	Office and Garage Trailers, Portable Toilets	$2,350	
	Generator/Fuel/Electrician	$4,000	
	Materials	$3,250	
Operating Costs (2-week Survey)			**$59,500**
	Supervisor	$12,000	
	On-site Analyst	$9,500	
	Vehicle Operator	$6,100	
	Field Technician	$7,500	
	Per Diem (4 Personnel x 14 days x $150)	$8,400	
	Rental Vehicles	$4,000	
	Equipment Repair	$12,000	
Analysis & Reporting			**$25,000**
Total Cost			**$161,600**

5.2 Cost Analysis

The relative productivity of the three vehicular MTADS sensor platforms is given in Table 10. The productivities listed are based on surveys of the two Standardized Test Sites. In a larger, more regular field, the productivity numbers would be uniformly higher but the trend would remain. Because of lower sampling rates, the EM systems must drive at half the speed of the magnetometer system. As was done for the demonstration for this program, discrimination-quality data requires either interleaved sampling (in the case of the GEM array) or surveying in two directions (in the case of the EM61 array). This requires another factor of two penalty on the EM systems for an overall productivity reduction of a factor of four.

Table 10. Measured Productivities of the Three MTADS Sensor Arrays.

Sensor Platform	Productivity (ha/hour)
Magnetometer Array	0.74
EM61 Array	0.30
GEM-3 Array	0.27

6. References

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

2. "Results of the MTADS Technology Demonstration #3, Jefferson Proving Ground, Madison, IN," J. R. McDonald, H. H. Nelson, R. Jeffries, and R. Robertson, NRL/PU/6110--99-375, January 1999.

3. "MTADS Unexploded Ordnance Operations at the Badlands Bombing Range, Pine Ridge Reservation, Cuny Table, SD," J. R. McDonald, H. H. Nelson, J. Neece, R. Robertson, and J. Jeffries, NRL/PU/6110--98-353, July 1997.

4. "Jefferson Proving Ground Technology Demonstration Program Summary," G. Robitaille, J. Adams, C. O'Donnell, and P. Burr, http://aec.army.mil/usaec/technology/jpgsummary.pdf.

5. "Frequency-Domain Electromagnetic Induction Sensors for the Multi-Sensor Towed Array Detection System," H. H. Nelson, B. Barrow, T. Bell, R. S. Jones, and B. SanFilipo, NRL/MR/6110—02-8650, November 2002.

6. "Design and Construction of the NRL Baseline Ordnance Classification Test Site at Blossom Point," Nelson, H.H., McDonald, J.R., Robertson, R., NRL/MR/6110--00-8437, March 2000.

7. "Electromagnetic Induction and Magnetic Sensor Fusion for Enhanced UXO Target Classification," H. H. Nelson and Bruce Barrow, NRL/PU/6110--00-423.

8. "Standardized UXO Technology Demonstration Site Blind Grid Scoring Record No. 27 (Naval Research Laboratory)," Overbay, L. and The Standardized UXO Technology Demonstration Site Scoring Committee, ATC-8740, January 2004.

9. "Standardized UXO Technology Demonstration Site Open Field Scoring Record No. 675 (Naval Research Laboratory)," Overbay, L., Robitaille, G., and The Standardized UXO Technology Demonstration Site Scoring Committee, ATC-8974, August 2005.

10. "Standardized UXO Technology Demonstration Site Blind Grid Scoring Record No. 213 (Naval Research Laboratory)," Overbay, L. and The Standardized UXO Technology Demonstration Site Scoring Committee, ATC-8836, January 2005.

11. "Standardized UXO Technology Demonstration Site Open Field Scoring Record No. 245 (Naval Research Laboratory)," Overbay, L., and The Standardized UXO Technology Demonstration Site Scoring Committee, ATC-8942, March 2005.

7. Points of Contact

ESTCP

Jeff Marqusee	Director, ESTCP	ESTCP 901 North Stuart Street, Suite 303 Arlington, VA 22203	Tel: 703-696-2120 Fax: 703-696-2114 jeffrey.marqusee@osd.mil
Anne Andrews	Program Manager, Munitions Management	ESTCP 901 North Stuart Street, Suite 303 Arlington, VA 22203	Tel: 703-696-3826 Fax: 703-696-2114 anne.andrews@osd.mil

Naval Research Laboratory

Herb Nelson	Project Lead	Code 6110 Bldg 207, Rm. 279A Naval Research Lab Washington, DC 20375	Tel: 202-767-3686 Fax: 202-404-8119 Cell: 202-215-4844 herb.nelson@nrl.navy.mil

Nova Research, Inc.

Dan Steinhurst	Data Analyst	1900 Elkin St. Suite 230 Alexandria, VA 22308	Tel: 202-767-3556 Fax: 202-404-8119 Cell: 703-850-5217 dan.steinhurst@nrl.navy.mil
Glenn Harbaugh	Field Technician	1900 Elkin St. Suite 230 Alexandria, VA 22308	Tel: 301-392-1702 Fax: 301-392-1702 Cell: 410-610-3506

AETC, Inc.

Tom Bell	Analysis Lead	1225 S. Clark Street Suite 800 Arlington, VA 22202	Tel: 703-413-0500 Fax: 703-413-0505 tbell@va.aetc.com
Bruce Barrow	Data Analyst	1225 S. Clark Street Suite 800 Arlington, VA 22202	Tel: 703-413-0500 Fax: 703-413-0505 bjb@va.aetc.com
Nagi Khadr	Data Analyst	1225 S. Clark Street Suite 800 Arlington, VA 22202	Tel: 217-531-9026 nagi@va.aetc.com

Geophex Ltd.

IJ Won	President	605 Mercury St. Raleigh, NC 27603-2343	Tel: 919 839-8515 Fax: 919- 839-8528 ijwon@geophex.com
Bill San Filipo	Data Analyst	605 Mercury St. Raleigh, NC 27603-2343	Tel: 919 839-8515 Fax: 919- 839-8528 sanfilipo@geophex.com

Appendix A. Summary of the APG Blind Grid Scoring Report

This and the following three appendices contain summaries of the official scoring reports for the GEM-3 array at the Standardized Test Sites.

Figure A1 shows the probability of detection for the response stage (P_d^{res}) and the discrimination stage (P_d^{disc}) versus their respective probability of false positive. Figure A2 shows both probabilities plotted against their respective probability of background alarm. Both figures use horizontal lines to illustrate the performance of the demonstrator at two demonstrator-specified points: at the system noise level for the response stage, representing the point below which targets are not considered detectable, and at the demonstrator's recommended threshold level for the discrimination stage, defining the subset of targets the demonstrator would recommend digging based on discrimination.

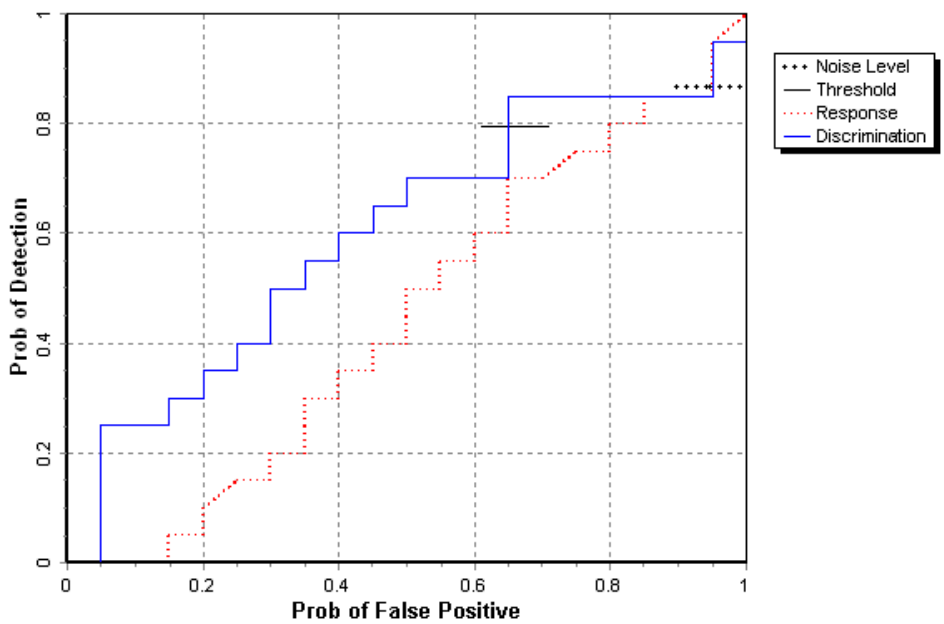

Figure A1 – APG Blind Grid probability of detection for response and discrimination stages versus their respective probability of false positive over all ordnance categories combined

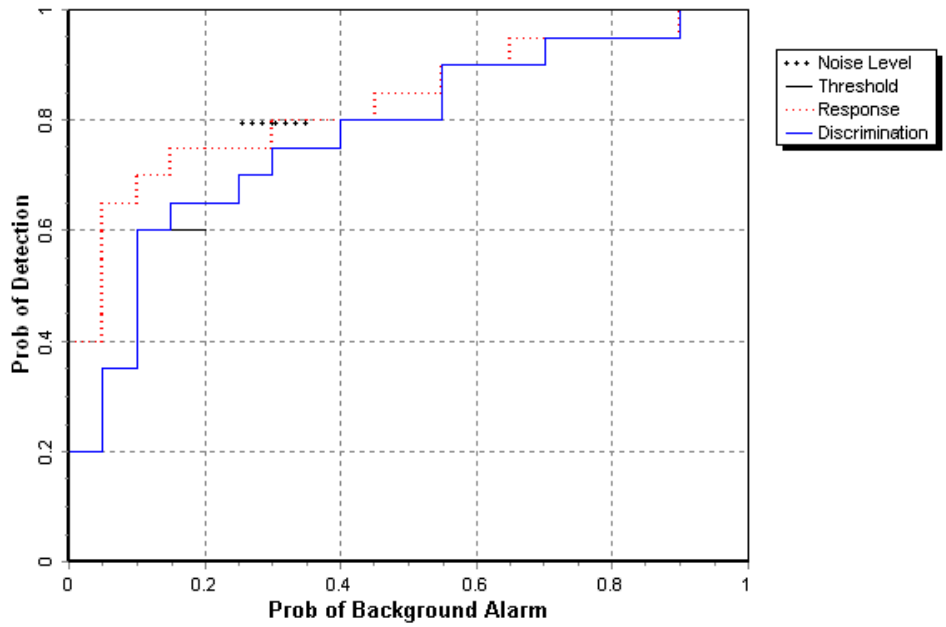

Figure A2 – APG Blind Grid probability of detection for response and discrimination stages versus their respective probability of background alarm over all ordnance categories combined

Results for the APG Blind Grid test, broken out by size, depth and nonstandard ordnance, are presented in Table A1. The results by size show how well the demonstrator did at detecting/discriminating ordnance of a certain caliber range. Depth is measured from the closest point of anomaly to the ground surface.

Table A1 Summary of APG Blind Grid Results

Metric	Overall	Standard	Nonstandard	By Size			By Depth, m		
				Small	Medium	Large	< 0.3	0.3 to <1	>= 1
RESPONSE STAGE									
P_d	0.85	0.90	0.85	1.00	0.75	0.80	1.00	0.95	0.00
P_d Low 90% Conf	0.81	0.80	0.73	0.91	0.61	0.55	0.95	0.88	0.00
P_{fp}	0.95	-	-	-	-	-	0.95	0.95	1.00
P_{fp} Low 90% Conf	0.90	-	-	-	-	-	0.89	0.85	0.63
P_{ba}	0.20	-	-	-	-	-	-	-	-
DISCRIMINATION STAGE									
P_d	0.80	0.80	0.80	0.95	0.60	0.70	0.90	0.90	0.00
P_d Low 90% Conf	0.73	0.69	0.69	0.88	0.48	0.45	0.82	0.80	0.00
P_{fp}	0.65	-	-	-	-	-	0.55	0.75	0.80
P_{fp} Low 90% Conf	0.59	-	-	-	-	-	0.47	0.63	0.42
P_{ba}	0.15	-	-	-	-	-	-	-	-

Response Stage Noise Level: 0.40
Recommended Discrimination Stage Threshold: 100.00

Efficiency and rejection rates are calculated to quantify the discrimination ability at specific points of interest on the ROC curve: (1) at the point where no decrease in P_d is suffered (i.e., the efficiency is by definition equal to one) and (2) at the operator selected threshold. These values are reported in Table A2

Table A2 APG Blind Grid Efficiency and Rejection Rates

	Efficiency (E)	**False Positive Rejection Rate**	**Background Alarm Rejection Rate**
At Operating Point	0.92	0.30	0.08
With No Loss of P_d	1.00	0.06	0.00

At the demonstrator's recommended setting, the ordnance items that were detected and correctly discriminated were further scored on whether their correct type could be identified (Table A3). Correct type examples include "20-mm projectile, 105-mm HEAT Projectile, and 2.75-in. Rocket". A list of the standard type declaration required for each ordnance item was provided to demonstrators prior to testing. For example, the standard type for the three example items are 20mmP, 105H, and 2.75in, respectively.

Table A3 Correct Type Classification of Targets Correctly Discriminated As UXO at the APG Blind Grid

Size	**% Correct**
Small	65.0
Medium	57.9
Large	71.4
Overall	63.6

The mean location error and standard deviations appear in Table A4. These calculations are based on average missed depth for ordnance correctly identified in the discrimination stage. Depths are measured from the closest point of the ordnance to the surface. For the Blind Grid, only depth errors are calculated, since (x, y) positions are known to be the centers of each grid square

Table A4 Mean Location Error and Standard Deviation (m) for the APG Blind Grid

	Mean	**Standard Deviation**
Depth	-0.06	0.20

Appendix B. Summary of the APG Open Field Scoring Report

Figure B1 shows the probability of detection for the response stage (P_d^{res}) and the discrimination stage (P_d^{disc}) versus their respective probability of false positive. Figure B2 shows both probabilities plotted against their respective probability of background alarm. Both figures use horizontal lines to illustrate the performance of the demonstrator at two demonstrator-specified points: at the system noise level for the response stage, representing the point below which targets are not considered detectable, and at the demonstrator's recommended threshold level for the discrimination stage, defining the subset of targets the demonstrator would recommend digging based on discrimination.

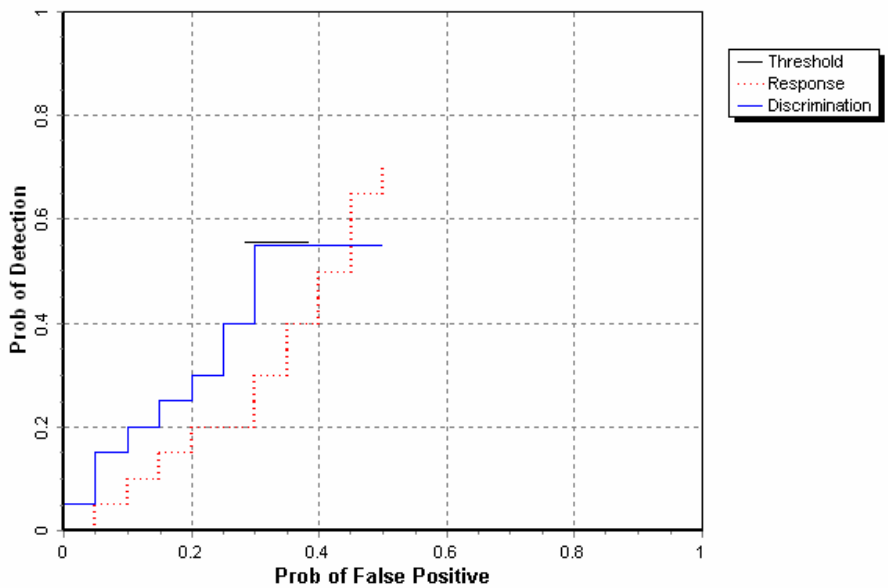

Figure B1 – APG Open Field probability of detection for response and discrimination stages versus their respective probability of false positive over all ordnance categories combined

Results for the APG Open Field test, broken out by size, depth and nonstandard ordnance, are presented in Table B1. The results by size show how well the demonstrator did at detecting/discriminating ordnance of a certain caliber range. Depth is measured from the closest point of anomaly to the ground surface.

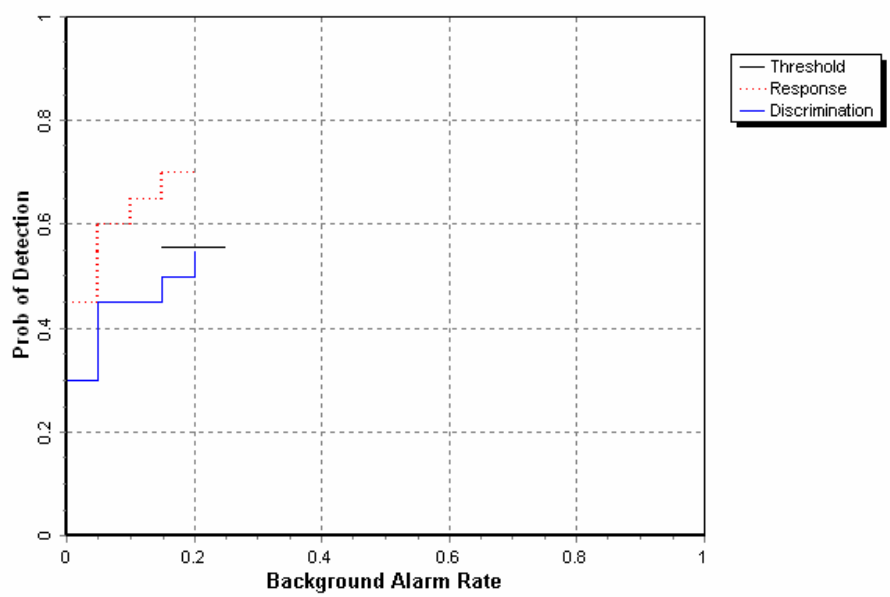

Figure B2 – APG Open Field probability of detection for response and discrimination stages versus their respective probability of background alarm over all ordnance categories combined

Table B1 Summary of APG Open Field Results

Metric	Overall	Standard	Nonstandard	By Size			By Depth, m		
				Small	Medium	Large	< 0.3	0.3 to <1	>= 1
RESPONSE STAGE									
P_d	0.70	0.75	0.65	0.70	0.70	0.80	0.80	0.65	0.50
P_d Low 90% Conf	0.68	0.71	0.60	0.65	0.62	0.71	0.77	0.61	0.44
P_d Upper 90% Conf	0.74	0.79	0.71	0.75	0.74	0.84	0.86	0.73	0.61
P_{fp}	0.50	-	-	-	-	-	0.45	0.55	0.70
P_{fp} Low 90% Conf	0.48	-	-	-	-	-	0.41	0.52	0.50
P_{fp} Upper 90% Conf	0.52	-	-	-	-	-	0.47	0.58	0.84
BAR	0.20	-	-	-	-	-	-	-	-
DISCRIMINATION STAGE									
P_d	0.55	0.55	0.55	0.55	0.55	0.60	0.60	0.55	0.45
P_d Low 90% Conf	0.52	0.51	0.49	0.51	0.47	0.49	0.54	0.51	0.34
P_d Upper 90% Conf	0.59	0.61	0.61	0.62	0.59	0.65	0.65	0.63	0.52
P_{fp}	0.35	-	-	-	-	-	0.30	0.35	0.45
P_{fp} Low 90% Conf	0.31	-	-	-	-	-	0.28	0.33	0.26
P_{fp} Upper 90% Conf	0.35	-	-	-	-	-	0.34	0.38	0.62
BAR	0.20	-	-	-	-	-	-	-	-

Response Stage Noise Level: 7.00
Recommended Discrimination Stage Threshold: 125.00

Efficiency and rejection rates are calculated to quantify the discrimination ability at specific points of interest on the ROC curve: (1) at the point where no decrease in P_d is suffered (i.e., the efficiency is by definition equal to one) and (2) at the operator selected threshold. These values are reported in Table B2

Table B2 APG Open Field Efficiency and Rejection Rates

	Efficiency (E)	**False Positive Rejection Rate**	**Background Alarm Rejection Rate**
At Operating Point	0.78	0.33	0.10
With No Loss of P_d	1.00	0.02	0.00

At the demonstrator's recommended setting, the ordnance items that were detected and correctly discriminated were further scored on whether their correct type could be identified (Table B3). Correct type examples include "20-mm projectile, 105-mm HEAT Projectile, and 2.75-in. Rocket". A list of the standard type declaration required for each ordnance item was provided to demonstrators prior to testing. For example, the standard type for the three example items are 20mmP, 105H, and 2.75in, respectively.

Table B3 Correct Type Classification of Targets Correctly Discriminated As UXO at the APG Open Field

Size	**% Correct**
Small	23.5
Medium	7.7
Large	33.3
Overall	20.3

The mean location error and standard deviations appear in Table B4. These calculations are based on average missed depth for ordnance correctly identified in the discrimination stage. Depths are measured from the closest point of the ordnance to the surface. For the Blind Grid, only depth errors are calculated, since (x, y) positions are known to be the centers of each grid square

Table B4 Mean Location Error and Standard Deviation (m) for the APG Open Field

	Mean	**Standard Deviation**
Northing	0.00	0.17
Easting	0.01	0.16
Depth	-0.07	0.43

Appendix C. Summary of the YPG Blind Grid Scoring Report

Figure C1 shows the probability of detection for the response stage (P_d^{res}) and the discrimination stage (P_d^{disc}) versus their respective probability of false positive. Figure C2 shows both probabilities plotted against their respective probability of background alarm. Both figures use horizontal lines to illustrate the performance of the demonstrator at two demonstrator-specified points: at the system noise level for the response stage, representing the point below which targets are not considered detectable, and at the demonstrator's recommended threshold level for the discrimination stage, defining the subset of targets the demonstrator would recommend digging based on discrimination.

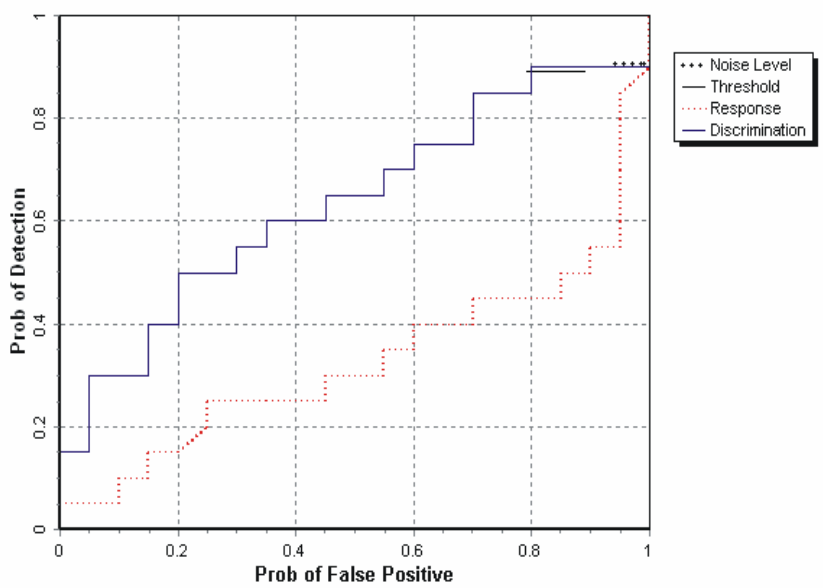

Figure C1 – YPG Blind Grid probability of detection for response and discrimination stages versus their respective probability of false positive over all ordnance categories combined

Results for the YPG Blind Grid test, broken out by size, depth and nonstandard ordnance, are presented in Table C1. The results by size show how well the demonstrator did at detecting/discriminating ordnance of a certain caliber range. Depth is measured from the closest point of anomaly to the ground surface.

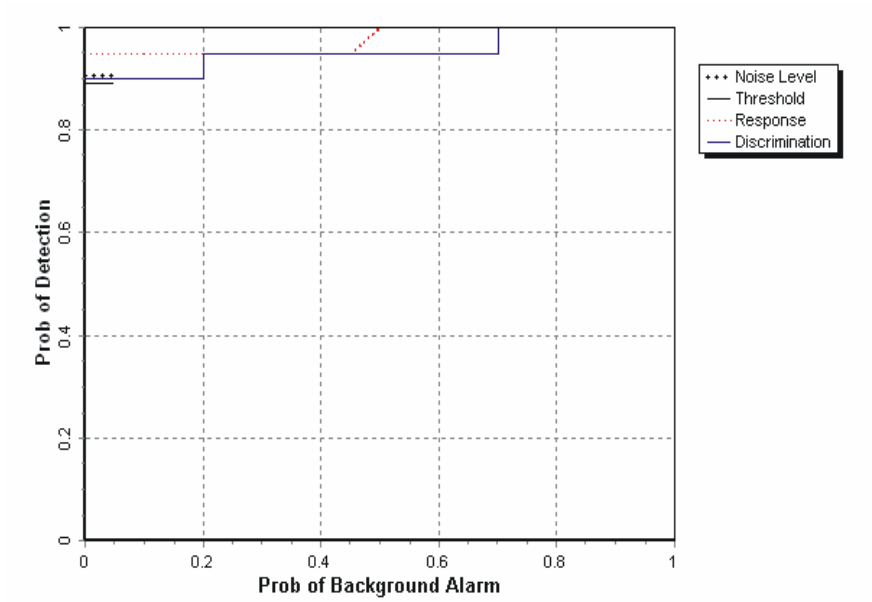

Figure C2 – YPG Blind Grid probability of detection for response and discrimination stages versus their respective probability of background alarm over all ordnance categories combined

Table C1 Summary of YPG Blind Grid Results

Metric	Overall	Standard	Nonstandard	By Size			By Depth, m		
				Small	Medium	Large	< 0.3	0.3 to <1	>= 1
RESPONSE STAGE									
P_d	0.90	0.90	0.90	0.95	0.85	0.95	1.00	0.90	0.30
P_d Low 90% Conf	0.85	0.83	0.78	0.86	0.686	0.75	0.95	0.79	0.08
P_{fp}	1.00	-	-	-	-	-	1.00	1.00	N/A
P_{fp} Low 90% Conf	0.97	-	-	-	-	-	0.96	0.92	-
P_{ba}	0.00	-	-	-	-	-	-	-	-
DISCRIMINATION STAGE									
P_d	0.90	0.90	0.85	0.95	0.80	0.95	1.00	0.90	0.30
P_d Low 90% Conf	0.83	0.83	0.74	0.86	0.63	0.75	0.91	0.79	0.08
P_{fp}	0.85	-	-	-	-	-	0.80	0.95	N/A
P_{fp} Low 90% Conf	0.79	-	-	-	-	-	0.74	0.87	-
P_{ba}	0.00	-	-	-	-	-	-	-	-

Response Stage Noise Level: 2.60
Recommended Discrimination Stage Threshold: 14.97

Efficiency and rejection rates are calculated to quantify the discrimination ability at specific points of interest on the ROC curve: (1) at the point where no decrease in P_d is suffered (i.e., the efficiency is by definition equal to one) and (2) at the operator selected threshold. These values are reported in Table C2

Table C2 YPG Blind Grid Efficiency and Rejection Rates

	Efficiency (E)	**False Positive Rejection Rate**	**Background Alarm Rejection Rate**
At Operating Point	0.99	0.15	Undefined
With No Loss of P_d	1.00	0.01	Undefined

At the demonstrator's recommended setting, the ordnance items that were detected and correctly discriminated were further scored on whether their correct type could be identified (Table C3). Correct type examples include "20-mm projectile, 105-mm HEAT Projectile, and 2.75-in. Rocket". A list of the standard type declaration required for each ordnance item was provided to demonstrators prior to testing. For example, the standard type for the three example items are 20mmP, 105H, and 2.75in, respectively.

Table C3 Correct Type Classification of Targets Correctly Discriminated As UXO at the YPG Blind Grid

Size	**% Correct**
Small	60.0
Medium	44.4
Large	15.4
Overall	47.0

The mean location error and standard deviations appear in Table C4. These calculations are based on average missed depth for ordnance correctly identified in the discrimination stage. Depths are measured from the closest point of the ordnance to the surface. For the Blind Grid, only depth errors are calculated, since (x, y) positions are known to be the centers of each grid square

Table C4 Mean Location Error and Standard Deviation (m) for the YPG Blind Grid

	Mean	**Standard Deviation**
Depth	-0.01	0.32

Appendix D. Summary of the YPG Open Field Scoring Report

Figure D1 shows the probability of detection for the response stage (P_d^{res}) and the discrimination stage (P_d^{disc}) versus their respective probability of false positive. Figure D2 shows both probabilities plotted against their respective probability of background alarm. Both figures use horizontal lines to illustrate the performance of the demonstrator at two demonstrator-specified points: at the system noise level for the response stage, representing the point below which targets are not considered detectable, and at the demonstrator's recommended threshold level for the discrimination stage, defining the subset of targets the demonstrator would recommend digging based on discrimination.

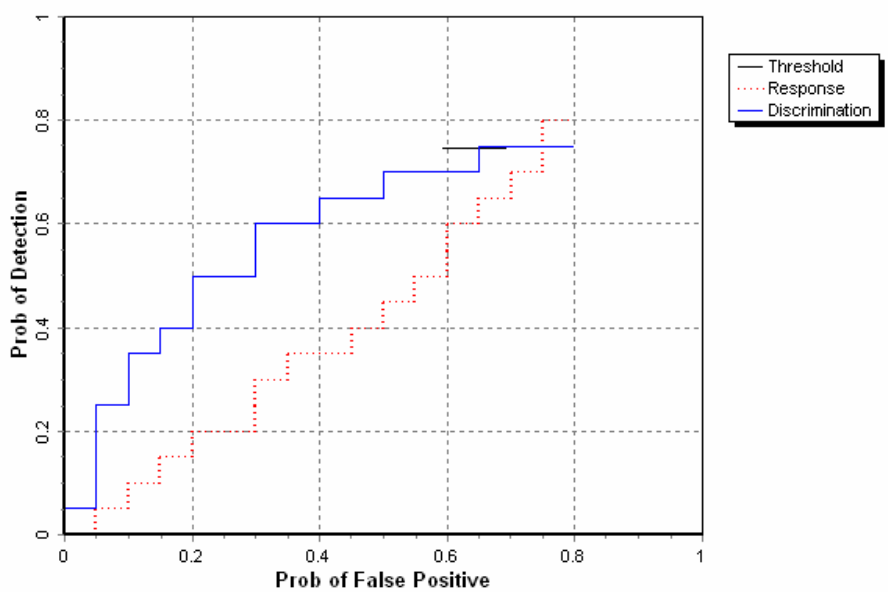

Figure D1 – YPG Open Field probability of detection for response and discrimination stages versus their respective probability of false positive over all ordnance categories combined

Results for the YPG Open Field test, broken out by size, depth and nonstandard ordnance, are presented in Table D1. The results by size show how well the demonstrator did at detecting/discriminating ordnance of a certain caliber range. Depth is measured from the closest point of anomaly to the ground surface.

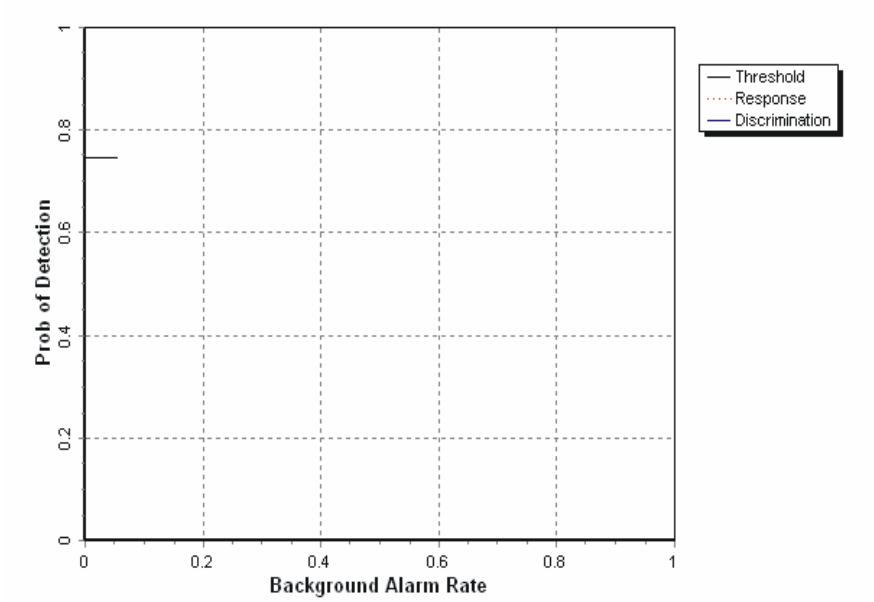

Figure D2 – YPG Open Field probability of detection for response and discrimination stages versus their respective probability of background alarm over all ordnance categories combined

Table D1 Summary of YPG Open Field Results

Metric	Overall	Standard	Nonstandard	By Size			By Depth, m		
				Small	Medium	Large	< 0.3	0.3 to <1	>= 1
RESPONSE STAGE									
P_d	0.80	0.80	0.80	0.80	0.75	0.90	0.80	0.80	0.50
P_d Low 90% Conf	0.76	0.75	0.75	0.74	0.69	0.83	0.79	0.76	0.38
P_{fp}	0.80	-	-	-	-	-	0.80	0.80	0.30
P_{fp} Low 90% Conf	0.76	-	-	-	-	-	0.76	0.76	0.12
BAR	0.00	-	-	-	-	-	-	-	-
DISCRIMINATION STAGE									
P_d	0.75	0.75	0.75	0.75	0.70	0.85	0.70	0.80	0.45
P_d Low 90% Conf	0.72	0.70	0.72	0.70	0.65	0.78	0.73	0.74	0.35
P_{fp}	0.65	-	-	-	-	-	0.60	0.75	0.30
P_{fp} Low 90% Conf	0.63	-	-	-	-	-	0.58	0.72	0.12
BAR	0.00	-	-	-	-	-	-	-	-

Response Stage Noise Level: 1.20
Recommended Discrimination Stage Threshold: 14.86

Efficiency and rejection rates are calculated to quantify the discrimination ability at specific points of interest on the ROC curve: (1) at the point where no decrease in P_d is suffered (i.e., the efficiency is by definition equal to one) and (2) at the operator selected threshold. These values are reported in Table D2

Table D2 YPG Open Field Efficiency and Rejection Rates

	Efficiency (E)	**False Positive Rejection Rate**	**Background Alarm Rejection Rate**
At Operating Point	0.95	0.17	0.36
With No Loss of P_d	1.00	0.02	1.00

At the demonstrator's recommended setting, the ordnance items that were detected and correctly discriminated were further scored on whether their correct type could be identified (Table D3). Correct type examples include "20-mm projectile, 105-mm HEAT Projectile, and 2.75-in. Rocket". A list of the standard type declaration required for each ordnance item was provided to demonstrators prior to testing. For example, the standard type for the three example items are 20mmP, 105H, and 2.75in, respectively.

Table D3 Correct Type Classification of Targets Correctly Discriminated As UXO at the YPG Open Field

Size	**% Correct**
Small	70.6
Medium	55.8
Large	39.5
Overall	60.4

The mean location error and standard deviations appear in Table D4. These calculations are based on average missed depth for ordnance correctly identified in the discrimination stage. Depths are measured from the closest point of the ordnance to the surface. For the Blind Grid, only depth errors are calculated, since (x, y) positions are known to be the centers of each grid square

Table D4 Mean Location Error and Standard Deviation (m) for the YPG Open Field

	Mean	**Standard Deviation**
Northing	0.00	0.08
Easting	0.00	0.10
Depth	-0.09	0.19