AD NO. _____
DTC PROJECT NO. 8-CO-160-UXO-021
REPORT NO. ATC-9149

STANDARDIZED

UXO TECHNOLOGY DEMONSTRATION SITE

BLIND GRID SCORING RECORD NO. 691

SITE LOCATION:
U.S. ARMY ABERDEEN PROVING GROUND

DEMONSTRATOR:
ARM GROUP INC.
1129 WEST GOVERNOR ROAD
P.O. BOX 797
HERSHEY, PA 17033

TECHNOLOGY TYPE/PLATFORM:
MINELAB F3/HAND HELD

PREPARED BY:
U.S. ARMY ABERDEEN TEST CENTER
ABERDEEN PROVING GROUND, MD 21005-5059

JANUARY 2006

Prepared for:
U.S. ARMY ENVIRONMENTAL CENTER
ABERDEEN PROVING GROUND, MD 21010-5401

U.S. ARMY DEVELOPMENTAL TEST COMMAND
ABERDEEN PROVING GROUND, MD 21005-5055

DISTRIBUTION UNLIMITED, JANUARY 2006.

DISTRIBUTION STATEMENT A
Approved for Public Release
Distribution Unlimited

DISPOSITION INSTRUCTIONS

Destroy this document when no longer needed. Do not return to the originator.

The use of trade names in this document does not constitute an official endorsement or approval of the use of such commercial hardware or software. This document may not be cited for purposes of advertisement.

REPORT DOCUMENTATION PAGE	Form Approved OMB No. 0704-0188

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.
PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
January 2006	Final	4 through 7 and 14 April 2005

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
STANDARDIZED UXO TECHNOLOGY DEMONSTRATION SITE BLIND GRID SCORING RECORD NO. 691 (ARM GROUP INC.)	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S) Overbay, Larry; Watts, Kimberly The Standardized UXO Technology Demonstration Site Scoring Committee	5d. PROJECT NUMBER 8-CO-160-UXO-021
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	8. PERFORMING ORGANIZATION REPORT NUMBER
Commander U.S. Army Aberdeen Test Center ATTN: CSTE-DTC-AT-SL-E Aberdeen Proving Ground, MD 21005-5059	ATC-9149
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Commander U.S. Army Environmental Center ATTN: SFIM-AEC-ATT Aberdeen Proving Ground, MD 21005-5401	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S) Same as Item 8

12. DISTRIBUTION/AVAILABILITY STATEMENT
Distribution unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
This scoring record documents the efforts of ARM Group Inc., to detect and discriminate inert unexploded ordnance (UXO) utilizing the APG Standardized UXO Technology Demonstration Site blind grid. Scoring Records have been coordinated by Larry Overbay and the Standardized UXO Technology Demonstration Site Scoring Committee. Organizations on the committee include, the U.S. Army Corps of Engineers, the Environmental Security Technology Certification Program, the Strategic Environmental Research and Development Program, the Institute for Defense Analysis, the U.S. Army Environmental Center, and the U.S. Army Aberdeen Test Center.

15. SUBJECT TERMS
ARM Group Inc., UXO Standardized Technology Demonstration Site Program, blind grid, EM GM-5 Minelab F3/hand held, MEC

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER *(Include area code)*
Unclassified	Unclassified	Unclassified	UL		

ACKNOWLEDGEMENTS

Authors:

Larry Overbay Jr.
Military Environmental Technology Demonstration Center (METDC)
U.S. Army Aberdeen Test Center (ATC)
U.S. Army Aberdeen Proving Ground (APG)

Rick Fling
Aberdeen Test and Support Services (ATSS)
Sverdrup Technology, Inc.
U.S. Army Aberdeen Proving Ground (APG)

Christina McClung
Aberdeen Data Services Team (ADST)
Logistics Engineering and Information Technology Company (Log.Sec)
U.S. Army Aberdeen Proving Ground (APG)

Contributor:

Kimberly Watts
U.S. Army Environmental Center (AEC)
U.S. Army Aberdeen Proving Ground (APG)

Matthew Banta
Aberdeen Test and Support Services (ATSS)
Sverdrup Technology, Inc.
U.S. Army Aberdeen Proving Ground (APG)

TABLE OF CONTENTS

		PAGE
	ACKNOWLEDGMENTS	i
	SECTION 1. GENERAL INFORMATION	
1.1	BACKGROUND	1
1.2	SCORING OBJECTIVES	1
	1.2.1 Scoring Methodology	1
	1.2.2 Scoring Factors	2
1.3	STANDARD AND NONSTANDARD INERT ORDNANCE TARGETS	3
	SECTION 2. DEMONSTRATION	
2.1	DEMONSTRATOR INFORMATION	5
	2.1.1 Demonstrator Point of Contact (POC) and Address	5
	2.1.2 System Description	5
	2.1.3 Data Processing Description	6
	2.1.4 Data Submission Format	7
	2.1.5 Demonstrator Quality Assurance (QA) and Quality Control (QC)	8
	2.1.6 Additional Records	8
2.2	APG SITE INFORMATION	9
	2.2.1 Location	9
	2.2.2 Soil Type	9
	2.2.3 Test Areas	9
	SECTION 3. FIELD DATA	
3.1	DATE OF FIELD ACTIVITIES	11
3.2	AREAS TESTED/NUMBER OF HOURS	11
3.3	TEST CONDITIONS	11
	3.3.1 Weather Conditions	11
	3.3.2 Field Conditions	11
	3.3.3 Soil Moisture	11
3.4	FIELD ACTIVITIES	12
	3.4.1 Setup/Mobilization	12
	3.4.2 Calibration	12
	3.4.3 Downtime Occasions	12
	3.4.4 Data Collection	12
	3.4.5 Demobilization	12
3.5	PROCESSING TIME	13
3.6	DEMONSTRATOR'S FIELD PERSONNEL	13
3.7	DEMONSTRATOR'S FIELD SURVEYING METHOD	13
3.8	SUMMARY OF DAILY LOGS	13

SECTION 4. TECHNICAL PERFORMANCE RESULTS

		PAGE
4.1	ROC CURVES USING ALL ORDNANCE CATEGORIES	15
4.2	ROC CURVES USING ORDNANCE LARGER THAN 20 MM	16
4.3	PERFORMANCE SUMMARIES	18
4.4	EFFICIENCY, REJECTION RATES, AND TYPE CLASSIFICATION	19
4.5	LOCATION ACCURACY	19

SECTION 5. ON-SITE LABOR COSTS

SECTION 6. COMPARISON OF RESULTS TO BLIND GRID DEMONSTRATION

SECTION 7. APPENDIXES

A	TERMS AND DEFINITIONS	A-1
B	DAILY WEATHER LOGS	B-1
C	SOIL MOISTURE	C-1
D	DAILY ACTIVITY LOGS	D-1
E	REFERENCES	E-1
F	ABBREVIATIONS	F-1
G	DISTRIBUTION LIST	G-1

SECTION 1. GENERAL INFORMATION

1.1 BACKGROUND

Technologies under development for the detection and discrimination of munitions and explosives of concern (MEC) – i.e. unexploded ordnance (UXO) and discarded military munitions (DMM) require testing so that their performance can be characterized. To that end, Standardized Test Sites have been developed at Aberdeen Proving Ground (APG), Maryland and U.S. Army Yuma Proving Ground (YPG), Arizona. These test sites provide a diversity of geology, climate, terrain, and weather as well as diversity in ordnance and clutter. Testing at these sites is independently administered and analyzed by the government for the purposes of characterizing technologies, tracking performance with system development, comparing performance of different systems, and comparing performance in different environments.

The Standardized UXO Technology Demonstration Site Program is a multi-agency program spearheaded by the U.S. Army Environmental Center (AEC). The U.S. Army Aberdeen Test Center (ATC) and the U.S. Army Corps of Engineers Engineering Research and Development Center (ERDC) provide programmatic support. The program is being funded and supported by the Environmental Security Technology Certification Program (ESTCP), the Strategic Environmental Research and Development Program (SERDP) and the Army Environmental Quality Technology Program (EQT).

1.2 SCORING OBJECTIVES

The objective in the Standardized UXO Technology Demonstration Site Program is to evaluate the detection and discrimination capabilities of a given technology under various field and soil conditions. Inert munitions and clutter items are positioned in various orientations and depths in the ground.

The evaluation objectives are as follows:

a. To determine detection and discrimination effectiveness under realistic scenarios that vary targets, geology, clutter, topography, and vegetation.

b. To determine cost, time, and manpower requirements to operate the technology.

c. To determine demonstrator's ability to analyze survey data in a timely manner and provide prioritized "Target Lists" with associated confidence levels.

d. To provide independent site management to enable the collection of high quality, ground-truth, geo-referenced data for post-demonstration analysis.

1.2.1 Scoring Methodology

a. The scoring of the demonstrator's performance is conducted in two stages. These two stages are termed the RESPONSE STAGE and DISCRIMINATION STAGE. For both stages, the probability of detection (P_d) and the false alarms are reported as receiver-operating

characteristic (ROC) curves. False alarms are divided into those anomalies that correspond to emplaced clutter items, measuring the probability of false positive (P_{fp}), and those that do not correspond to any known item, termed background alarms.

b. The RESPONSE STAGE scoring evaluates the ability of the system to detect emplaced targets without regard to ability to discriminate ordnance from other anomalies. For the blind grid RESPONSE STAGE, the demonstrator provides the scoring committee with a target response from each and every grid square along with a noise level below which target responses are deemed insufficient to warrant further investigation. This list is generated with minimal processing and, since a value is provided for every grid square, will include signals both above and below the system noise level.

c. The DISCRIMINATION STAGE evaluates the demonstrator's ability to correctly identify ordnance as such and to reject clutter. For the blind grid DISCRIMINATION STAGE, the demonstrator provides the scoring committee with the output of the algorithms applied in the discrimination-stage processing for each grid square. The values in this list are prioritized based on the demonstrator's determination that a grid square is likely to contain ordnance. Thus, higher output values are indicative of higher confidence that an ordnance item is present at the specified location. For digital signal processing, priority ranking is based on algorithm output. For other discrimination approaches, priority ranking is based on human (subjective) judgment. The demonstrator also specifies the threshold in the prioritized ranking that provides optimum performance, (i.e. that is expected to retain all detected ordnance and rejects the maximum amount of clutter).

d. The demonstrator is also scored on EFFICIENCY and REJECTION RATIO, which measures the effectiveness of the discrimination stage processing. The goal of discrimination is to retain the greatest number of ordnance detections from the anomaly list, while rejecting the maximum number of anomalies arising from non-ordnance items. EFFICIENCY measures the fraction of detected ordnance retained after discrimination, while the REJECTION RATIO measures the fraction of false alarms rejected. Both measures are defined relative to performance at the demonstrator-supplied level below which all responses are considered noise, i.e., the maximum ordnance detectable by the sensor and its accompanying false positive rate or background alarm rate.

e. All scoring factors are generated utilizing the Standardized UXO Probability and Plot Program, version 3.1.1.

1.2.2 Scoring Factors

Factors to be measured and evaluated as part of this demonstration include:

a. Response Stage ROC curves:

(1) Probability of Detection (P_d^{res}).

(2) Probability of False Positive (P_{fp}^{res}).

(3) Background Alarm Rate (BAR^{res}) or Probability of Background Alarm (P_{BA}^{res}).

b. Discrimination Stage ROC curves:

(1) Probability of Detection (P_d^{disc}).

(2) Probability of False Positive (P_{fp}^{disc}).

(3) Background Alarm Rate (BAR^{disc}) or Probability of Background Alarm (P_{BA}^{disc}).

c. Metrics:

(1) Efficiency (E).

(2) False Positive Rejection Rate (R_{fp}).

(3) Background Alarm Rejection Rate (R_{BA}).

d. Other:

(1) Probability of Detection by Size and Depth.

(2) Classification by type (i.e., 20-mm, 40-mm, 105-mm, etc.).

(3) Location accuracy.

(4) Equipment setup, calibration time and corresponding man-hour requirements.

(5) Survey time and corresponding man-hour requirements.

(6) Reacquisition/resurvey time and man-hour requirements (if any).

(7) Downtime due to system malfunctions and maintenance requirements.

1.3 STANDARD AND NONSTANDARD INERT ORDNANCE TARGETS

The standard and nonstandard ordnance items emplaced in the test areas are listed in Table 1. Standardized targets are members of a set of specific ordnance items that have identical properties to all other items in the set (caliber, configuration, size, weight, aspect ratio, material, filler, magnetic remanence, and nomenclature). Nonstandard targets are ordnance items having properties that differ from those in the set of standardized targets.

TABLE 1. INERT ORDNANCE TARGETS

Standard Type	Nonstandard (NS)
20-mm Projectile M55	20-mm Projectile M55
	20-mm Projectile M97
40-mm Grenades M385	40-mm Grenades M385
40-mm Projectile MKII Bodies	40-mm Projectile M813
BDU-28 Submunition	
BLU-26 Submunition	
M42 Submunition	
57-mm Projectile APC M86	
60-mm Mortar M49A3	60-mm Mortar (JPG)
	60-mm Mortar M49
2.75-inch Rocket M230	2.75-inch Rocket M230
	2.75-inch Rocket XM229
MK 118 ROCKEYE	
81-mm Mortar M374	81-mm Mortar (JPG)
	81-mm Mortar M374
105-mm HEAT Rounds M456	
105-mm Projectile M60	105-mm Projectile M60
155-mm Projectile M483A1	155-mm Projectile M483A
	500-lb Bomb
	M75 Submunition

JPG = Jefferson Proving Ground
HEAT = high-explosive antitank

SECTION 2. DEMONSTRATION

2.1 DEMONSTRATOR INFORMATION

2.1.1 Demonstrator Point of Contact (POC) and Address

POC: Jeffrey Leberfinger, P.G.
717-533-8600

Address: ARM Group Inc.
1129 West Governor Road, (P.O. Box 797)
Hershey, PA 17033

2.1.2 System Description (provided by demonstrator)

The Minelab Explorer II metal detector is designed as a metal detector for amateur prospectors. It is currently being used on a number of projects as an unexploded ordnance (UXO) detector because it incorporates features that are considered advantageous to detect and discriminate between UXO-related material targets and non-UXO-related material targets. The system can be considered a hybrid of better known PEMI and FD systems in that it transmits a CW waveform based on 28 frequencies ranging between 1.5 and 100 kHz but uses time domain gating techniques to demodulate the signal and derive the target response.

The main feature that has attracted the interest of the UXO community is the real-time target discrimination capability: a two-dimensional map that plots a target response according to its electrical conductivity (on the horizontal axis of the screen) and a purely ferrous response (vertical axis). Operators quickly learn to recognize well-known objects such as bottle tops, coins, nails, etc., based on their on-screen location. Another feature of the Explorer II is its use of sophisticated filtering algorithms to process the target responses, specifically to enhance the signal-to-noise ratio of a detected target against the background response due to the soil or rock. Finally, the Explorer II must be used with a continuous sweeping action across the ground, and the sweeping speed can aid in target discrimination.

The Minelab F3 metal detector (fig. 1) is the latest product to be developed by Minelab for landmine detection; eventually, the F3 will supersede the F1A4. The F3 incorporates all of the features of Minelab's patented and well-known Multi Period Sensing technology that has been very successful at eliminating magnetic soil responses. The main difference is that the transmitted waveform in the F3 is bipolar. This type of waveform was developed to produce an instrument that would not set off certain classes of landmines that respond to conventional unipolar electromagnetic induction.

Similar to the F1B2 and F1A4, the F3 provides a two-channel output, available through a serial RS232 interface at the rate of 100 Hz, or alternatively can be used for detection using the audio tone. The most significant difference between these instruments is that the F3 incorporates more digital logic circuitry and microprocessing power. The F3 also has a waveform consisting of a positive long pulse and a negative shorter pulse period with significantly higher amplitude that ensures both pulses have an equal area, resulting in a net zero effect on magnetic mines. The two different pulse lengths are the specific innovations that provide the ability to null out the geologic noise in real time.

A significant improvement in the F3 over previous versions is that it operates in a direct current (DC) mode all the time, and it does not suffer from drift in the same way as the F1 series. Operators find this feature particularly useful if they are accustomed to the F1 series, which required them to repeatedly sweep the coil across the ground at a fast speed in order to detect targets. The coil can be moved as fast or as slow as the operator requires since it does not have an alternating current (AC) filter that can attenuate a response with a slow-moving coil as is the case of the F1 series of detectors.

Figure 1. Demonstrator's system, Minelab F3/hand held.

2.1.3 Data Processing Description (provided by demonstrator)

Sweeping the APG Calibration Grid.

The APG calibration grid, which is approximately 0.30 acres, will be surveyed with the Explorer II and F3 in a manner that follows a typical detect/flag/interrogate style survey which involves a sweeping motion at a constant height above ground surface. Both instruments use an

audio tone to indicate the presence of a target, and in such surveys, the point where the tone is maximum is usually the location where a flag is placed for further detailed detection, discrimination, and investigation. The operator will be instructed to take advantage of knowing exactly where each of the targets are located, in order to ensure the best possible signals are obtained during the training. The marked locations will be swept with the sensor coil passing directly over the flag with a left to right motion while the data being is being recorded to utilize the special target classification features that indicate the amplitudes and polarities while inferring the amount of conductive and ferrous material in the target. The axes position of the response can be stored as a rough signature representative of these tow parameter values. The stored parameters and features will be used to classify seeded items (at unknown locations to the operator) in the Blind Grid and Active Response Site.

Sweeping the APG blind grid.

The APG blind grid, which is approximately 0.48 acres, will also be surveyed with the Explorer II and F3 in a similar manner (a sweeping motion) as the calibration grid with three major exceptions. The first exception is that lanes will be marked (every 1 meter along the grid) so the operator can simulate real-world detect/flag operations by sweeping lanes until the entire grid is effectively surveyed. The second exception is that no targets will be pre-marked with pin flags. The third exception is that the operator will now use the recorded training information gathered from the calibration grid to ascertain whether each item is UXO-related or non-UXO-related material through real-time discrimination in the blind grid. The operator will walk the lanes while using the same sweeping motion, only stopping for audible sounds to mark the peak anomaly location and further discriminate/classify the item. The marked peak anomaly location will be surveyed using Global Positioning System (GPS) and all operator descriptions regarding each anomaly location will be documented. Discrimination effectiveness will be tabulated as to whether the detector/operator system is able to: (1) classify a difference between UXO and non-UXO related material and (2) classify the actual UXO item. As a footnote, a non-response could be due to: the metallic seeded item response is beyond the range of the detector, the metallic seeded item response is outside the range of the parameters for the seeded items in the Calibration Grid, or no metal is present at that location. The reasons for a non-response will not be known to the operator but could be determined in post-project analysis and discussions.

Sweeping Portions of the APG Active Response Site.

A portion (approximately 2.0 acres) of the APG Active Response Site will be surveyed in a similar manner (a sweeping motion) as the blind grid.

2.1.4 Data Submission Format

Data were submitted for scoring in accordance with data submission protocols outlined in the Standardized UXO Technology Demonstration Site Handbook. These submitted data are not included in this report in order to protect ground truth information.

2.1.5 Demonstrator Quality Assurance (QA) and Quality Control (QC) (provided by demonstrator)

Each of the two detector teams will consist of two people, one to operate the system, the other to record information on a data sheet as the survey progresses and to keep an eye on the detector operator to ensure good technique is maintained at all times. Prior to arrival at APG, detector teams will be instructed on strict protocols relating to field technique and documentation of anomalies. At the beginning of each day, detectors will be tested by confirming their ability to detect selected items on the calibration grid using the audio tone. The items will be specifically chosen as being representative of targets close to the limit of detection. As each anomaly is detected, it will be assigned a unique ID number which will be entered into the field notes as well as the computer record relating to the stored anomaly profile. Each team will be assigned a number range to ensure the whole site is covered with unique numbers. Operators will be trained to pay particular attention to note taking. In order to ensure detectors will be swept across every inch of ground, the operators have been given plenty of time to perform their assigned tasks so they do not feel they have to rush through the lanes. Operators will be fully briefed on the importance of their good technique and practices in this demonstration because the results will be used as a baseline for future testing of the F3 and Explorer II detectors.

2.1.6 Additional Records

The following record(s) by this vendor can be accessed via the Internet as MicroSoft Word documents at www.uxotestsites.org.

2.2 APG SITE INFORMATION

2.2.1 Location

The APG Standardized Test Site is located within a secured range area of the Aberdeen Area of APG. The Aberdeen Area of APG is located approximately 30 miles northeast of Baltimore at the northern end of the Chesapeake Bay. The Standardized Test Site encompasses 17 acres of upland and lowland flats, woods, and wetlands.

2.2.2 Soil Type

According to the soils survey conducted for the entire area of APG in 1998, the test site consists primarily of Elkton Series type soil (ref 2). The Elkton Series consists of very deep, slowly permeable, poorly drained soils. These soils formed in silty aeolin sediments and the underlying loamy alluvial and marine sediments. They are on upland and lowland flats and in depressions of the Mid-Atlantic Coastal Plain. Slopes range from 0 to 2 percent.

ERDC conducted a site-specific analysis in May of 2002 (ref 3). The results basically matched the soil survey mentioned above. Seventy percent of the samples taken were classified as silty loam. The majority (77 percent) of the soil samples had a measured water content between 15- and 30-percent with the water content decreasing slightly with depth.

For more details concerning the soil properties at the APG test site, go to www.uxotestsites.org on the web to view the entire soils description report.

2.2.3 Test Areas

A description of the test site areas at APG is included in Table 2.

TABLE 2. TEST SITE AREAS

Area	Description
Calibration Grid	Contains 14 standard ordnance items buried in six positions at various angles and depths to allow demonstrator equipment calibration.
Blind Grid	Contains 400 grid cells in a 0.2-hectare (0.5 acre) site. The center of each grid cell contains ordnance, clutter or nothing.

(Page 10 Blank)

SECTION 3. FIELD DATA

3.1 DATE OF FIELD ACTIVITIES (4 through 7 and 14 April 2005)

3.2 AREAS TESTED/NUMBER OF HOURS

Areas tested and total number of hours operated at each site are summarized in Table 3.

TABLE 3. AREAS TESTED AND NUMBER OF HOURS

Area	Number of Hours
Calibration Lanes	12.16
Blind Grid	17.25

3.3 TEST CONDITIONS

3.3.1 Weather Conditions

An APG weather station located approximately one mile west of the test site was used to record average temperature and precipitation on a half hour basis for each day of operation. The temperatures listed in Table 4 represent the average temperature during field operations from 0700 to 1700 hours while precipitation data represents a daily total amount of rainfall. Hourly weather logs used to generate this summary are provided in Appendix B.

TABLE 4. TEMPERATURE/PRECIPITATION DATA SUMMARY

Date, 2005	Average Temperature, °F	Total Daily Precipitation, in.
04 April	58.00	0.00
05 April	59.63	0.00
06 April	68.93	0.00
07 April	70.50	0.00
14 April	56.02	0.00

3.3.2 Field Conditions

The weather was cool. The blind grid was wet, and standing areas of water were present due to rain prior to testing.

3.3.3 Soil Moisture

Three soil probes were placed at various locations within the site to capture soil moisture data: Calibration, Mogul, Open Field, and Wooded areas. Measurements were collected in percent moisture and were taken twice daily (morning and afternoon) from five different soil depths (1 to 6 in., 6 to 12 in., 12 to 24 in., 24 to 36 in., and 36 to 48 in.) from each probe. Soil moisture logs are included in Appendix C.

3.4 FIELD ACTIVITIES

3.4.1 Setup/Mobilization

These activities included initial mobilization and daily equipment preparation and break down. A two-person crew took 1 hour and 5 minutes to perform the initial setup and mobilization. There was 1 hour and 5 minutes of daily equipment preparation, and end of the day equipment break down lasted 20 minutes.

3.4.2 Calibration

ARM Group spent a total of 12 hours and 10 minutes in the calibration lanes, of which 7 hours and 45 minutes was spent collecting data.

3.4.3 Downtime Occasions

Occasions of downtime are grouped into five categories: equipment/data checks or equipment maintenance, equipment failure and repair, weather, Demonstration Site issues, or breaks/lunch. All downtime is included for the purposes of calculating labor costs (section 5) except for downtime due to Demonstration Site issues. Demonstration Site issues, while noted in the Daily Log, are considered non-chargeable downtime for the purposes of calculating labor costs and are not discussed. Breaks and lunches are discussed in this section and billed to the total Site Survey area.

3.4.3.1 Equipment/data checks, maintenance. Equipment data checks and maintenance activities accounted for 5 minutes of site usage time. These activities included changing out batteries and routine data checks to ensure the data was being properly recorded/collected. ARM Group spent an additional 2 hours and 45 minutes for breaks and lunches.

3.4.3.2 Equipment failure or repair. No time was needed to resolve equipment failures that occurred while surveying the blind grid.

3.4.3.3 Weather. No weather delays occurred during the survey.

3.4.4 Data Collection

ARM Group spent a total time of 17 hours and 15 minutes in the blind grid area, 13 hours of which was spent collecting data.

3.4.5 Demobilization

The ARM Group survey crew went on to conduct a demonstration of the active site. Therefore, demobilization did not occur until 14 April 2005. On that day, it took the crew 30 minutes to break down and pack up their equipment.

3.5 PROCESSING TIME

ARM Group submitted the raw data from the demonstration activities on the last day of the demonstration, as required. The scoring submittal data was also provided on 5 May 2005.

3.6 DEMONSTRATOR'S FIELD PERSONNEL

Brian Brunette, Recorder/GPS Operator: Project Geophysicist
Alex Mussio, Recorder/GPS Operator: Geophysicist
Christopher Parker, Explorer II Operator: UXO Technician
Terry Foot, F3 Operator: UXO Technician

3.7 DEMONSTRATOR'S FIELD SURVEYING METHOD

ARM Group surveyed the blind grid in a north to south direction with a 1-meter line spacing through the middle of each grid lane.

3.8 SUMMARY OF DAILY LOGS

Daily logs capture all field activities during this demonstration and are located in Appendix D. Activities pertinent to this specific demonstration are indicated in highlighted text.

(Page 14 Blank)

SECTION 4. TECHNICAL PERFORMANCE RESULTS

4.1 ROC CURVES USING ALL ORDNANCE CATEGORIES

Figure 2 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 3 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. Note that all points have been rounded to protect the ground truth.

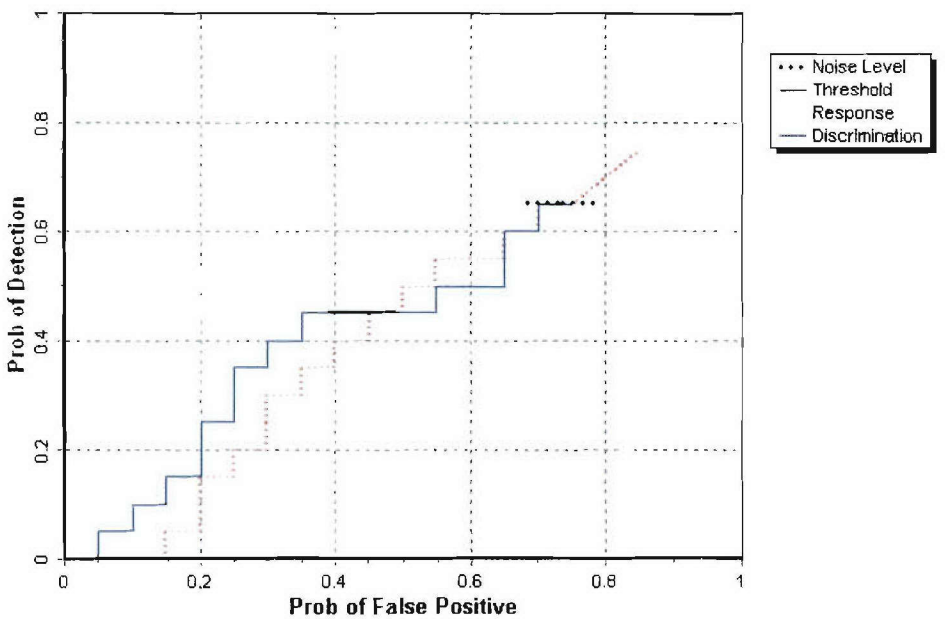

Figure 2. F3/hand held blind grid probability of detection for response and discrimination stages versus their respective probability of false positive over all ordnance categories combined.

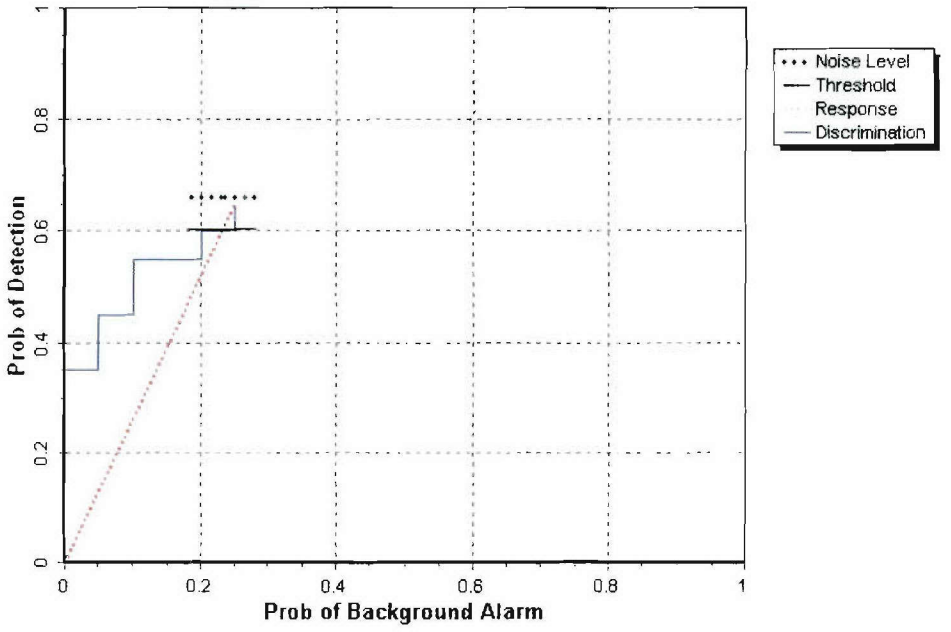

Figure 3. Explorer II/hand held blind grid probability of detection for response and discrimination stages versus their respective probability of background alarm over all ordnance categories combined.

4.2 ROC CURVES USING ORDNANCE LARGER THAN 20 MM

Figure 4 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 when only targets larger than 20 mm are scored. Figure 5 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. Note that all points have been rounded to protect the ground truth.

Note: The response stages shown below contain binomial distributions which are indicated by a straight line within the associated figures.

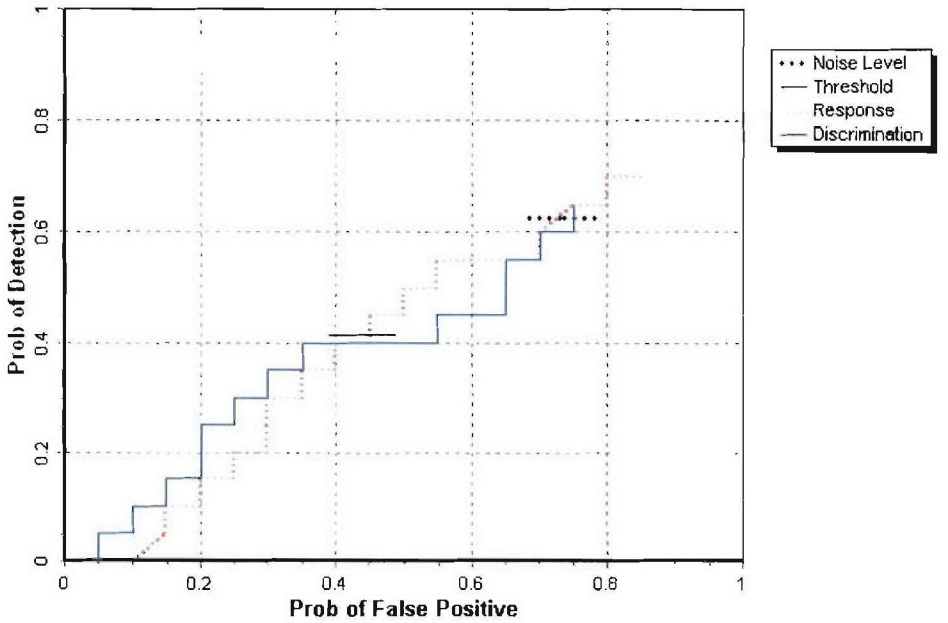

Figure 4. F3/hand held blind grid probability of detection for response and discrimination stages versus their respective probability of false positive for all ordnance larger than 20 mm.

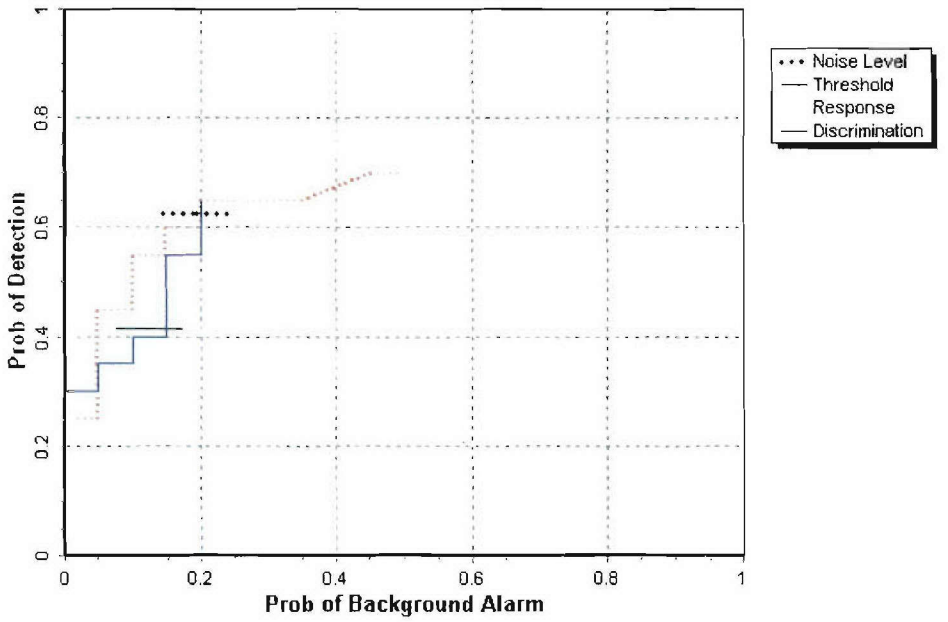

Figure 5. F3/hand held blind grid probability of detection for response and discrimination stages versus their respective probabilities of background alarm for all ordnance larger than 20 mm.

4.3 PERFORMANCE SUMMARIES

Results for the Blind Grid test broken out by size, depth and nonstandard ordnance are presented in Table 5 (for cost results, see section 5). Results by size and depth include both standard and nonstandard ordnance. The results by size show how well the demonstrator did at detecting/discriminating ordnance of a certain caliber range (see app A for size definitions). The results are relative to the number of ordnance items emplaced. Depth is measured from the geometric center of anomalies.

The RESPONSE STAGE results are derived from the list of anomalies above the demonstrator-provided noise level. The results for the DISCRIMINATION STAGE are derived from the demonstrator's recommended threshold for optimizing UXO field cleanup by minimizing false digs and maximizing ordnance recovery. The lower 90 percent confidence limit on probability of detection and P_{fp} was calculated assuming that the number of detections and false positives are binomially distributed random variables. All results in Table 5 have been rounded to protect the ground truth. However, lower confidence limits were calculated using actual results.

TABLE 5. SUMMARY OF BLIND GRID RESULTS FOR THE F3/HAND HELD

Metric	Overall	Standard	Nonstandard	By Size			By Depth, m		
				Small	Medium	Large	< 0.3	0.3 to <1	>= 1
RESPONSE STAGE									
P_d	0.65	0.60	0.70	0.75	0.60	0.40	0.95	0.55	0.25
P_d Low 90% Conf	0.58	0.51	0.59	0.66	0.45	0.19	0.86	0.42	0.11
P_d Upper 90% Conf	0.72	0.70	0.81	0.85	0.69	0.65	0.98	0.66	0.44
P_{fp}	0.75	-	-	-	-	-	0.95	0.60	0.35
P_{fp} Low 90% Conf	0.67	-	-	-	-	-	0.88	0.50	0.09
P_d Upper 90% Conf	0.79	-	-	-	-	-	0.99	0.69	0.67
P_{ba}	0.20	-	-	-	-	-	-	-	-
DISCRIMINATION STAGE									
P_d	0.45	0.40	0.50	0.65	0.25	0.20	0.85	0.25	0.05
P_d Low 90% Conf	0.38	0.32	0.39	0.54	0.17	0.05	0.72	0.16	0.01
P_d Upper 90% Conf	0.53	0.51	0.63	0.75	0.40	0.45	0.91	0.38	0.22
P_{fp}	0.45	-	-	-	-	-	0.60	0.35	0.00
P_{fp} Low 90% Conf	0.37	-	-	-	-	-	0.51	0.25	0.00
P_d Upper 90% Conf	0.51	-	-	-	-	-	0.72	0.44	0.32
P_{ba}	0.15	-	-	-	-	-	-	-	-

Response Stage Noise Level: 10.50
Recommended Discrimination Stage Threshold: 61.50

Note: The recommended discrimination stage threshold values are provided by the demonstrator.

4.4 EFFICIENCY, REJECTION RATES, AND TYPE CLASSIFICATION

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 6.

TABLE 6. EFFICIENCY AND REJECTION RATES

	Efficiency (E)	False Positive Rejection Rate	Background Alarm Rejection Rate
At Operating Point	0.70	0.40	0.36
With No Loss of P_d	1.00	0.00	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 7). Correct type examples include "20-mm projectile, 105-mm HEAT Projectile, and 2.75-inch 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 7. CORRECT TYPE CLASSIFICATION OF TARGETS CORRECTLY DISCRIMINATED AS UXO

Size	Percentage Correct
Small	57.1
Medium	22.2
Large	0.0
Overall	46.2

4.5 LOCATION ACCURACY

The mean location error and standard deviations appear in Table 8. 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 8. MEAN LOCATION ERROR AND STANDARD DEVIATION (M)

	Mean	Standard Deviation
Depth	-0.26	0.24

SECTION 5. ON-SITE LABOR COSTS

A standardized estimate for labor costs associated with this effort was calculated as follows: the first person at the test site was designated "supervisor", the second person was designated "data analyst", and the third and following personnel were considered "field support". Standardized hourly labor rates were charged by title: supervisor at $95.00/hour, data analyst at $57.00/hour, and field support at $28.50/hour.

Government representatives monitored on-site activity. All on-site activities were grouped into one of ten categories: initial setup/mobilization, daily setup/stop, calibration, collecting data, downtime due to break/lunch, downtime due to equipment failure, downtime due to equipment/data checks or maintenance, downtime due to weather, downtime due to demonstration site issue, or demobilization. See Appendix D for the daily activity log. See section 3.4 for a summary of field activities.

The standardized cost estimate associated with the labor needed to perform the field activities is presented in Table 9. Note that calibration time includes time spent in the Calibration Lanes as well as field calibrations. "Site survey time" includes daily setup/stop time, collecting data, breaks/lunch, downtime due to equipment/data checks or maintenance, downtime due to failure, and downtime due to weather.

TABLE 9. ON-SITE LABOR COSTS

	No. People	Hourly Wage	Hours	Cost
Initial Setup				
Supervisor	1	$95.00	1.08	$102.60
Data Analyst	1	57.00	1.08	61.56
Field Support	0	28.50	1.08	0.00
SubTotal				**$164.16**
Calibration				
Supervisor	1	$95.00	12.16	$1155.20
Data Analyst	1	57.00	12.16	693.12
Field Support	0	28.50	12.16	0.00
SubTotal				**$1,848.32**
Site Survey				
Supervisor	1	$95.00	17.25	$1638.75
Data Analyst	1	57.00	17.25	983.25
Field Support	0	28.50	17.25	0.00
SubTotal				**$2,622.00**

See notes at end of table.

TABLE 9 (CONT'D)

	No. People	Hourly Wage	Hours	Cost
Demobilization				
Supervisor	1	$95.00	0.50	$47.50
Data Analyst	1	57.00	0.50	28.50
Field Support	0	28.50	0.50	0.00
Subtotal				**$76.00**
Total				**$4,710.48**

Notes: Calibration time includes time spent in the Calibration Lanes as well as calibration before each data run.

Site Survey time includes daily setup/stop time, collecting data, breaks/lunch, downtime due to system maintenance, failure, and weather.

SECTION 6. COMPARISON OF RESULTS TO DATE

No comparisons to date.

(Page 24 Blank)

SECTION 7. APPENDIXES

APPENDIX A. TERMS AND DEFINITIONS

GENERAL DEFINITIONS

Anomaly: Location of a system response deemed to warrant further investigation by the demonstrator for consideration as an emplaced ordnance item.

Detection: An anomaly location that is within R_{halo} of an emplaced ordnance item.

Munitions and Explosives Of Concern (MEC): Specific categories of military munitions that may pose unique explosive safety risks, including UXO as defined in 10 USC 101(e)(5), DMM as defined in 10 USC 2710(e)(2) and/or munitions constituents (e.g. TNT, RDX) as defined in 10 USC 2710(e)(3) that are present in high enough concentrations to pose an explosive hazard.

Emplaced Ordnance: An ordnance item buried by the government at a specified location in the test site.

Emplaced Clutter: A clutter item (i.e., non-ordnance item) buried by the government at a specified location in the test site.

R_{halo}: A pre-determined radius about the periphery of an emplaced item (clutter or ordnance) within which a location identified by the demonstrator as being of interest is considered to be a response from that item. If multiple declarations lie within R_{halo} of any item (clutter or ordnance), the declaration with the highest signal output within the R_{halo} will be utilized. For the purpose of this program, a circular halo 0.5 meters in radius will be placed around the center of the object for all clutter and ordnance items less than 0.6 meters in length. When ordnance items are longer than 0.6 meters, the halo becomes an ellipse where the minor axis remains 1 meter and the major axis is equal to the length of the ordnance plus 1 meter.

Small Ordnance: Caliber of ordnance less than or equal to 40 mm (includes 20-mm projectile, 40-mm projectile, submunitions BLU-26, BLU-63, and M42).

Medium Ordnance: Caliber of ordnance greater than 40 mm and less than or equal to 81 mm (includes 57-mm projectile, 60-mm mortar, 2.75 in. Rocket, MK118 Rockeye, 81-mm mortar).

Large Ordnance: Caliber of ordnance greater than 81 mm (includes 105-mm HEAT, 105-mm projectile, 155-mm projectile, 500-pound bomb).

Shallow: Items buried less than 0.3 meter below ground surface.

Medium: Items buried greater than or equal to 0.3 meter and less than 1 meter below ground surface.

Deep: Items buried greater than or equal to 1 meter below ground surface.

Response Stage Noise Level: The level that represents the point below which anomalies are not considered detectable. Demonstrators are required to provide the recommended noise level for the Blind Grid test area.

Discrimination Stage Threshold: The demonstrator selected threshold level that they believe provides optimum performance of the system by retaining all detectable ordnance and rejecting the maximum amount of clutter. This level defines the subset of anomalies the demonstrator would recommend digging based on discrimination.

Binomially Distributed Random Variable: A random variable of the type which has only two possible outcomes, say success and failure, is repeated for n independent trials with the probability p of success and the probability 1-p of failure being the same for each trial. The number of successes x observed in the n trials is an estimate of p and is considered to be a binomially distributed random variable.

RESPONSE AND DISCRIMINATION STAGE DATA

The scoring of the demonstrator's performance is conducted in two stages. These two stages are termed the RESPONSE STAGE and DISCRIMINATION STAGE. For both stages, the probability of detection (P_d) and the false alarms are reported as receiver operating characteristic (ROC) curves. False alarms are divided into those anomalies that correspond to emplaced clutter items, measuring the probability of false positive (P_{fp}) and those that do not correspond to any known item, termed background alarms.

The RESPONSE STAGE scoring evaluates the ability of the system to detect emplaced targets without regard to ability to discriminate ordnance from other anomalies. For the RESPONSE STAGE, the demonstrator provides the scoring committee with the location and signal strength of all anomalies that the demonstrator has deemed sufficient to warrant further investigation and/or processing as potential emplaced ordnance items. This list is generated with minimal processing (e.g., this list will include all signals above the system noise threshold). As such, it represents the most inclusive list of anomalies.

The DISCRIMINATION STAGE evaluates the demonstrator's ability to correctly identify ordnance as such, and to reject clutter. For the same locations as in the RESPONSE STAGE anomaly list, the DISCRIMINATION STAGE list contains the output of the algorithms applied in the discrimination-stage processing. This list is prioritized based on the demonstrator's determination that an anomaly location is likely to contain ordnance. Thus, higher output values are indicative of higher confidence that an ordnance item is present at the specified location. For electronic signal processing, priority ranking is based on algorithm output. For other systems, priority ranking is based on human judgment. The demonstrator also selects the threshold that the demonstrator believes will provide "optimum" system performance, (i.e., that retains all the detected ordnance and rejects the maximum amount of clutter).

Note: The two lists provided by the demonstrator contain identical numbers of potential target locations. They differ only in the priority ranking of the declarations.

RESPONSE STAGE DEFINITIONS

Response Stage Probability of Detection (P_d^{res}): P_d^{res} = (No. of response-stage detections)/ (No. of emplaced ordnance in the test site).

Response Stage False Positive (fp^{res}): An anomaly location that is within R_{halo} of an emplaced clutter item.

Response Stage Probability of False Positive (P_{fp}^{res}): P_{fp}^{res} = (No. of response-stage false positives)/(No. of emplaced clutter items).

Response Stage Background Alarm (ba^{res}): An anomaly in a blind grid cell that contains neither emplaced ordnance nor an emplaced clutter item. An anomaly location in the open field or scenarios that is outside R_{halo} of any emplaced ordnance or emplaced clutter item.

Response Stage Probability of Background Alarm (P_{ba}^{res}): Blind Grid only: P_{ba}^{res} = (No. of response-stage background alarms)/(No. of empty grid locations).

Response Stage Background Alarm Rate (BAR^{res}): Open Field only: BAR^{res} = (No. of response-stage background alarms)/(arbitrary constant).

Note that the quantities P_d^{res}, P_{fp}^{res}, P_{ba}^{res}, and BAR^{res} are functions of t^{res}, the threshold applied to the response-stage signal strength. These quantities can therefore be written as $P_d^{res}(t^{res})$, $P_{fp}^{res}(t^{res})$, $P_{ba}^{res}(t^{res})$, and $BAR^{res}(t^{res})$.

DISCRIMINATION STAGE DEFINITIONS

Discrimination: The application of a signal processing algorithm or human judgment to response-stage data that discriminates ordnance from clutter. Discrimination should identify anomalies that the demonstrator has high confidence correspond to ordnance, as well as those that the demonstrator has high confidence correspond to non-ordnance or background returns. The former should be ranked with highest priority and the latter with lowest.

Discrimination Stage Probability of Detection (P_d^{disc}): P_d^{disc} = (No. of discrimination-stage detections)/(No. of emplaced ordnance in the test site).

Discrimination Stage False Positive (fp^{disc}): An anomaly location that is within R_{halo} of an emplaced clutter item.

Discrimination Stage Probability of False Positive (P_{fp}^{disc}): P_{fp}^{disc} = (No. of discrimination stage false positives)/(No. of emplaced clutter items).

Discrimination Stage Background Alarm (ba^{disc}): An anomaly in a blind grid cell that contains neither emplaced ordnance nor an emplaced clutter item. An anomaly location in the open field or scenarios that is outside R_{halo} of any emplaced ordnance or emplaced clutter item.

Discrimination Stage Probability of Background Alarm (P_{ba}^{disc}): P_{ba}^{disc} = (No. of discrimination-stage background alarms)/(No. of empty grid locations).

Discrimination Stage Background Alarm Rate (BAR^{disc}): BAR^{disc} = (No. of discrimination-stage background alarms)/(arbitrary constant).

Note that the quantities P_d^{disc}, P_{fp}^{disc}, P_{ba}^{disc}, and BAR^{disc} are functions of t^{disc}, the threshold applied to the discrimination-stage signal strength. These quantities can therefore be written as $P_d^{disc}(t^{disc})$, $P_{fp}^{disc}(t^{disc})$, $P_{ba}^{disc}(t^{disc})$, and $BAR^{disc}(t^{disc})$.

RECEIVER-OPERATING CHARACERISTIC (ROC) CURVES

ROC curves at both the response and discrimination stages can be constructed based on the above definitions. The ROC curves plot the relationship between P_d versus P_{fp} and P_d versus BAR or P_{ba} as the threshold applied to the signal strength is varied from its minimum (t_{min}) to its maximum (t_{max}) value.[1] Figure A-1 shows how P_d versus P_{fp} and P_d versus BAR are combined into ROC curves. Note that the "res" and "disc" superscripts have been suppressed from all the variables for clarity.

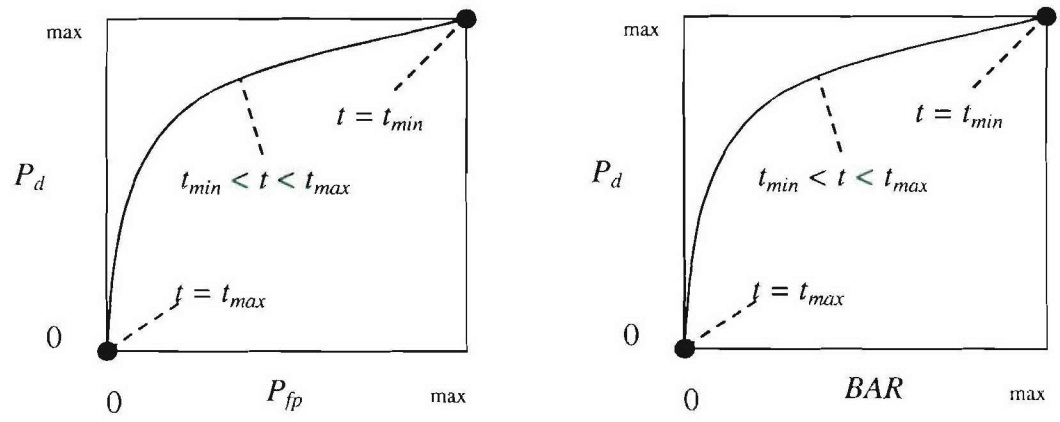

Figure A-1. ROC curves for open field testing. Each curve applies to both the response and discrimination stages.

[1]Strictly speaking, ROC curves plot the P_d versus P_{ba} over a pre-determined and fixed number of detection opportunities (some of the opportunities are located over ordnance and others are located over clutter or blank spots). In an open field scenario, each system suppresses its signal strength reports until some bare-minimum signal response is received by the system. Consequently, the open field ROC curves do not have information from low signal-output locations, and, furthermore, different contractors report their signals over a different set of locations on the ground. These ROC curves are thus not true to the strict definition of ROC curves as defined in textbooks on detection theory. Note, however, that the ROC curves obtained in the Blind Grid test sites are true ROC curves.

METRICS TO CHARACTERIZE THE DISCRIMINATION STAGE

The demonstrator is also scored on efficiency and rejection ratio, which measure the effectiveness of the discrimination stage processing. The goal of discrimination is to retain the greatest number of ordnance detections from the anomaly list, while rejecting the maximum number of anomalies arising from non-ordnance items. The efficiency measures the amount of detected ordnance retained by the discrimination, while the rejection ratio measures the fraction of false alarms rejected. Both measures are defined relative to the entire response list, i.e., the maximum ordnance detectable by the sensor and its accompanying false positive rate or background alarm rate.

Efficiency (E): $E = P_d^{disc}(t^{disc})/P_d^{res}(t_{min}^{res})$; Measures (at a threshold of interest), the degree to which the maximum theoretical detection performance of the sensor system (as determined by the response stage tmin) is preserved after application of discrimination techniques. Efficiency is a number between 0 and 1. An efficiency of 1 implies that all of the ordnance initially detected in the response stage was retained at the specified threshold in the discrimination stage, t^{disc}.

False Positive Rejection Rate (R_{fp}): $R_{fp} = 1 - [P_{fp}^{disc}(t^{disc})/P_{fp}^{res}(t_{min}^{res})]$; Measures (at a threshold of interest), the degree to which the sensor system's false positive performance is improved over the maximum false positive performance (as determined by the response stage tmin). The rejection rate is a number between 0 and 1. A rejection rate of 1 implies that all emplaced clutter initially detected in the response stage were correctly rejected at the specified threshold in the discrimination stage.

Background Alarm Rejection Rate (R_{ba}):

Blind Grid: $R_{ba} = 1 - [P_{ba}^{disc}(t^{disc})/P_{ba}^{res}(t_{min}^{res})]$.
Open Field: $R_{ba} = 1 - [BAR^{disc}(t^{disc})/BAR^{res}(t_{min}^{res})])$.

Measures the degree to which the discrimination stage correctly rejects background alarms initially detected in the response stage. The rejection rate is a number between 0 and 1. A rejection rate of 1 implies that all background alarms initially detected in the response stage were rejected at the specified threshold in the discrimination stage.

CHI-SQUARE COMPARISON EXPLANATION:

The Chi-square test for differences in probabilities (or 2 x 2 contingency table) is used to analyze two samples drawn from two different populations to see if both populations have the same or different proportions of elements in a certain category. More specifically, two random samples are drawn, one from each population, to test the null hypothesis that the probability of event A (some specified event) is the same for both populations (ref 3).

A 2 x 2 contingency table is used in the Standardized UXO Technology Demonstration Site Program to determine if there is reason to believe that the proportion of ordnance correctly detected/discriminated by demonstrator X's system is significantly degraded by the more challenging terrain feature introduced. The test statistic of the 2 x 2 contingency table is the

Chi-square distribution with one degree of freedom. Since an association between the more challenging terrain feature and relatively degraded performance is sought, a one-sided test is performed. A significance level of 0.05 is chosen which sets a critical decision limit of 2.71 from the Chi-square distribution with one degree of freedom. It is a critical decision limit because if the test statistic calculated from the data exceeds this value, the two proportions tested will be considered significantly different. If the test statistic calculated from the data is less than this value, the two proportions tested will be considered not significantly different.

An exception must be applied when either a 0 or 100 percent success rate occurs in the sample data. The Chi-square test cannot be used in these instances. Instead, Fischer's test is used and the critical decision limit for one-sided tests is the chosen significance level, which in this case is 0.05. With Fischer's test, if the test statistic is less than the critical value, the proportions are considered to be significantly different.

Standardized UXO Technology Demonstration Site examples, where blind grid results are compared to those from the open field and open field results are compared to those from one of the scenarios, follow. It should be noted that a significant result does not prove a cause and effect relationship exists between the two populations of interest; however, it does serve as a tool to indicate that one data set has experienced a degradation in system performance at a large enough level than can be accounted for merely by chance or random variation. Note also that a result that is not significant indicates that there is not enough evidence to declare that anything more than chance or random variation within the same population is at work between the two data sets being compared.

Demonstrator X achieves the following overall results after surveying each of the three progressively more difficult areas using the same system (results indicate the number of ordnance detected divided by the number of ordnance emplaced):

	Blind Grid	Open Field	Moguls
P_d^{res}	100/100 = 1.0	8/10 = .80	20/33 = .61
P_d^{disc}	80/100 = 0.80	6/10 = .60	8/33 = .24

P_d^{res}: BLIND GRID versus OPEN FIELD. Using the example data above to compare probabilities of detection in the response stage, all 100 ordnance out of 100 emplaced ordnance items were detected in the blind grid while 8 ordnance out of 10 emplaced were detected in the open field. Fischer's test must be used since a 100 percent success rate occurs in the data. Fischer's test uses the four input values to calculate a test statistic of 0.0075 that is compared against the critical value of 0.05. Since the test statistic is less than the critical value, the smaller response stage detection rate (0.80) is considered to be significantly less at the 0.05 level of significance. While a significant result does not prove a cause and effect relationship exists between the change in survey area and degradation in performance, it does indicate that the detection ability of demonstrator X's system seems to have been degraded in the open field relative to results from the blind grid using the same system.

P_d^{disc}: BLIND GRID versus OPEN FIELD. Using the example data above to compare probabilities of detection in the discrimination stage, 80 out of 100 emplaced ordnance items were correctly discriminated as ordnance in blind grid testing while 6 ordnance out of 10 emplaced were correctly discriminated as such in open field-testing. Those four values are used to calculate a test statistic of 1.12. Since the test statistic is less than the critical value of 2.71, the two discrimination stage detection rates are considered to be not significantly different at the 0.05 level of significance.

P_d^{res}: OPEN FIELD versus MOGULS. Using the example data above to compare probabilities of detection in the response stage, 8 out of 10 and 20 out of 33 are used to calculate a test statistic of 0.56. Since the test statistic is less than the critical value of 2.71, the two response stage detection rates are considered to be not significantly different at the 0.05 level of significance.

P_d^{disc}: OPEN FIELD versus MOGULS. Using the example data above to compare probabilities of detection in the discrimination stage, 6 out of 10 and 8 out of 33 are used to calculate a test statistic of 2.98. Since the test statistic is greater than the critical value of 2.71, the smaller discrimination stage detection rate is considered to be significantly less at the 0.05 level of significance. While a significant result does not prove a cause and effect relationship exists between the change in survey area and degradation in performance, it does indicate that the ability of demonstrator X to correctly discriminate seems to have been degraded by the mogul terrain relative to results from the flat open field using the same system.

(Page A-8 Blank)

APPENDIX B. DAILY WEATHER LOGS

Date & Time	Average Temperature, °F	Total Precipitation, in.
04/04/2005 0700	44.8	0
04/04/2005 0800	48.5	0
04/04/2005 0900	52.4	0
04/04/2005 1000	55.7	0
04/04/2005 1100	57.8	0
04/04/2005 1200	60.1	0
04/04/2005 1300	62	0
04/04/2005 1400	63.9	0
04/04/2005 1500	64.5	0
04/04/2005 1600	64.5	0
04/04/2005 1700	63.8	0
04/05/2005 0700	35.4	0
04/05/2005 0800	45.8	0
04/05/2005 0900	56.1	0
04/05/2005 1000	59	0
04/05/2005 1100	61.2	0
04/05/2005 1200	63.3	0
04/05/2005 1300	64.5	0
04/05/2005 1400	65.5	0
04/05/2005 1500	67.9	0
04/05/2005 1600	68.6	0
04/05/2005 1700	68.7	0
04/06/2005 0700	47.3	0
04/06/2005 0800	55.1	0
04/06/2005 0900	59.8	0
04/06/2005 1000	63.4	0
04/06/2005 1100	68.6	0
04/06/2005 1200	71.7	0
04/06/2005 1300	72.5	0
04/06/2005 1400	76.9	0
04/06/2005 1500	79.6	0
04/06/2005 1600	81.7	0
04/06/2005 1700	81.6	0
04/07/2005 0700	63.1	0
04/07/2005 0800	65	0
04/07/2005 0900	66.3	0
04/07/2005 1000	67.9	0
04/07/2005 1100	71	0
04/07/2005 1200	72	0
04/07/2005 1300	73.7	0
04/07/2005 1400	73.4	0
04/07/2005 1500	74.1	0
04/07/2005 1600	75.5	0
04/07/2005 1700	73.5	0

Date & Time	Average Temperature, °F	Total Precipitation, in.
04/08/2005 0700	59.9	0
04/08/2005 0800	59.1	0
04/08/2005 0900	57.6	0
04/08/2005 1000	58.3	0
04/08/2005 1100	59.2	0
04/08/2005 1200	59.6	0
04/08/2005 1300	61.6	0
04/08/2005 1400	61.4	0
04/08/2005 1500	62.1	0
04/08/2005 1600	63.8	0
04/08/2005 1700	62.9	0
04/09/2005 0700	48.4	0
04/09/2005 0800	52.4	0
04/09/2005 0900	55.2	0
04/09/2005 1000	57.7	0
04/09/2005 1100	59.4	0
04/09/2005 1200	61	0
04/09/2005 1300	62.4	0
04/09/2005 1400	63.6	0
04/09/2005 1500	64.1	0
04/09/2005 1600	64.3	0
04/09/2005 1700	64.2	0
04/10/2005 0700	40.1	0
04/10/2005 0800	49.8	0
04/10/2005 0900	54.7	0
04/10/2005 1000	59.2	0
04/10/2005 1100	63.6	0
04/10/2005 1200	67.6	0
04/10/2005 1300	68.6	0
04/10/2005 1400	70.3	0
04/10/2005 1500	71.7	0
04/10/2005 1600	72.2	0
04/10/2005 1700	72.1	0
04/11/2005 0700	55.5	0
04/11/2005 0800	57.3	0
04/11/2005 0900	60.1	0
04/11/2005 1000	61.7	0
04/11/2005 1100	62.7	0
04/11/2005 1200	63.4	0
04/11/2005 1300	65.2	0
04/11/2005 1400	66.5	0
04/11/2005 1500	67.1	0
04/11/2005 1600	67.2	0
04/11/2005 1700	64.5	0

Date & Time	Average Temperature, °F	Total Precipitation, in.
04/12/2005 0700	42.6	0
04/12/2005 0800	44.8	0
04/12/2005 0900	46.8	0
04/12/2005 1000	48.5	0
04/12/2005 1100	50.2	0
04/12/2005 1200	52.3	0
04/12/2005 1300	54.7	0
04/12/2005 1400	55.9	0
04/12/2005 1500	56	0
04/12/2005 1600	56.4	0
04/12/2005 1700	54.8	0
04/13/2005 0700	41.3	0
04/13/2005 0800	47.2	0
04/13/2005 0900	50	0
04/13/2005 1000	53.2	0
04/13/2005 1100	56.7	0
04/13/2005 1200	58.7	0
04/13/2005 1300	59.9	0
04/13/2005 1400	61.4	0
04/13/2005 1500	61.6	0
04/13/2005 1600	61.9	0
04/13/2005 1700	61.8	0
04/14/2005 0700	47.2	0
04/14/2005 0800	51.1	0
04/14/2005 0900	54.8	0
04/14/2005 1000	58.2	0
04/14/2005 1100	60.1	0
04/14/2005 1200	61.4	0
04/14/2005 1300	63.7	0
04/14/2005 1400	65.8	0
04/14/2005 1500	66.8	0
04/14/2005 1600	67.3	0
04/14/2005 1700	67	0
04/15/2005 0700	44	0
04/15/2005 0800	46.3	0
04/15/2005 0900	48.5	0
04/15/2005 1000	50.8	0
04/15/2005 1100	53	0
04/15/2005 1200	54.4	0
04/15/2005 1300	55.9	0
04/15/2005 1400	56.6	0
04/15/2005 1500	57.4	0
04/15/2005 1600	57.3	0
04/15/2005 1700	56.9	0
04/15/2005 1800	55.9	0
04/15/2005 1900	53.8	0
04/15/2005 2000	49.6	0
04/15/2005 2100	48.7	0
04/15/2005 2200	46.5	0
04/15/2005 2300	45.2	0

(Page B-4 Blank)

APPENDIX C. SOIL MOISTURE

Date: 4/04/2005
Times: 0800 through 1600

Probe Location	Layer, in.	AM Reading, %	PM Reading, %
Wet Area	0 to 6	NA	NA
	6 to 12	NA	NA
	12 to 24	NA	NA
	24 to 36	NA	NA
	36 to 48	NA	NA
Wooded Area	0 to 6	NA	NA
	6 to 12	NA	NA
	12 to 24	NA	NA
	24 to 36	NA	NA
	36 to 48	NA	NA
Open Area	0 to 6	NA	NA
	6 to 12	NA	NA
	12 to 24	NA	NA
	24 to 36	NA	NA
	36 to 48	NA	NA
Calibration Lanes	0 to 6	6.3	6.6
	6 to 12	38.8	36.9
	12 to 24	50.6	50.8
	24 to 36	45.5	44.9
	36 to 48	40.5	40.8
Blind Grid/Moguls	0 to 6	NA	NA
	6 to 12	NA	NA
	12 to 24	NA	NA
	24 to 36	NA	NA
	36 to 48	NA	NA

Date: 4/05/2005
Times: 0800 through 1600

Probe Location	Layer, in.	AM Reading, %	PM Reading, %
Wet Area	0 to 6	NA	NA
	6 to 12	NA	NA
	12 to 24	NA	NA
	24 to 36	NA	NA
	36 to 48	NA	NA
Wooded Area	0 to 6	NA	NA
	6 to 12	NA	NA
	12 to 24	NA	NA
	24 to 36	NA	NA
	36 to 48	NA	NA
Open Area	0 to 6	NA	NA
	6 to 12	NA	NA
	12 to 24	NA	NA
	24 to 36	NA	NA
	36 to 48	NA	NA
Calibration Lanes	0 to 6	6.2	6.1
	6 to 12	38.2	37.6
	12 to 24	50.6	50.4
	24 to 36	45.0	44.9
	36 to 48	40.1	40.0
Blind Grid/Moguls	0 to 6	3.9	3.9
	6 to 12	24.5	24.7
	12 to 24	38.0	38.1
	24 to 36	35.5	35.0
	36 to 48	39.7	40.3

Date: 4/06/2005
Times: 0800 through 1500

Probe Location	Layer, in.	AM Reading, %	PM Reading, %
Wet Area	0 to 6	NA	NA
	6 to 12	NA	NA
	12 to 24	NA	NA
	24 to 36	NA	NA
	36 to 48	NA	NA
Wooded Area	0 to 6	NA	NA
	6 to 12	NA	NA
	12 to 24	NA	NA
	24 to 36	NA	NA
	36 to 48	NA	NA
Open Area	0 to 6	NA	NA
	6 to 12	NA	NA
	12 to 24	NA	NA
	24 to 36	NA	NA
	36 to 48	NA	NA
Calibration Lanes	0 to 6	6.2	6.1
	6 to 12	38.2	37.6
	12 to 24	50.6	50.4
	24 to 36	45.0	44.9
	36 to 48	40.1	40.0
Blind Grid/Moguls	0 to 6	3.8	3.5
	6 to 12	24.1	24.3
	12 to 24	38.3	38.0
	24 to 36	35.2	35.0
	36 to 48	39.9	39.7

Date: 4/07/2005
Times: 0730 through 1500

Probe Location	Layer, in.	AM Reading, %	PM Reading, %
Wet Area	0 to 6	NA	NA
	6 to 12	NA	NA
	12 to 24	NA	NA
	24 to 36	NA	NA
	36 to 48	NA	NA
Wooded Area	0 to 6	NA	NA
	6 to 12	NA	NA
	12 to 24	NA	NA
	24 to 36	NA	NA
	36 to 48	NA	NA
Open Area	0 to 6	NA	NA
	6 to 12	NA	NA
	12 to 24	NA	NA
	24 to 36	NA	NA
	36 to 48	NA	NA
Calibration Lanes	0 to 6	6.0	6.0
	6 to 12	37.4	37.2
	12 to 24	50.4	50.6
	24 to 36	44.5	44.3
	36 to 48	39.6	39.5
Blind Grid/Moguls	0 to 6	3.8	3.7
	6 to 12	24.0	24.0
	12 to 24	37.8	38.0
	24 to 36	35.0	35.3
	36 to 48	39.6	39.8

Date	No. of People	Area-Tested	Status Start Time	Status Stop Time	Duration min.	Operational Status	Operational Status - Comments	Track Method	Track Method=Other Explain	Pattern	Field Conditions
4/04/2005	2	CALIBRATION LANES	855	1000	65	INITIAL SETUP		GPS	NA	LINEAR	SUNNY, MUDDY
4/04/2005	2	CALIBRATION LANES	1000	1150	110	COLLECTING DATA		GPS	NA	LINEAR	SUNNY, MUDDY
4/04/2005	2	CALIBRATION LANES	1150	1335	105	BREAK/LUNCH		GPS	NA	LINEAR	SUNNY, MUDDY
4/04/2005	2	CALIBRATION LANES	1335	1340	5	DOWNTIME DUE TO EQUIP MAINT/CHECK	CHANGE BATTERY	GPS	NA	LINEAR	SUNNY, MUDDY
4/04/2005	2	CALIBRATION LANES	1340	1630	170	COLLECTING DATA		GPS	NA	LINEAR	SUNNY, MUDDY
4/04/2005	2	CALIBRATION LANES	1630	1650	20	DAILY START, STOP	EQUIPMENT BREAKDOWN	GPS	NA	LINEAR	SUNNY, MUDDY
04/05/2005	2	CALIBRATION LANES	840	900	20	DAILY START, STOP		GPS	NA	LINEAR	SUNNY, MUDDY
04/05/2005	2	CALIBRATION LANES	900	1050	110	COLLECTING DATA		GPS	NA	LINEAR	SUNNY, MUDDY
04/05/2005	2	CALIBRATION LANES	1050	1115	25	BREAK/LUNCH		GPS	NA	LINEAR	SUNNY, MUDDY
04/05/2005	2	CALIBRATION LANES	1115	1200	45	COLLECTING DATA		GPS	NA	LINEAR	SUNNY, MUDDY
04/05/2005	2	CALIBRATION LANES	1200	1330	90	BREAK/LUNCH		GPS	NA	LINEAR	SUNNY, MUDDY
04/05/2005	2	CALIBRATION LANES	1330	1400	30	COLLECTING DATA		GPS	NA	LINEAR	SUNNY, MUDDY
04/05/2005	2	BLIND GRID	1400	1435	35	DAILY START, STOP	SET UP TAPES	GPS	NA	LINEAR	SUNNY, MUDDY
04/05/2005	2	BLIND GRID	1435	1630	115	COLLECTING DATA		GPS	NA	LINEAR	SUNNY, MUDDY
04/06/2005	2	BLIND GRID	1630	1640	10	DAILY START, STOP	EQUIPMENT BREAKDOWN	GPS	NA	LINEAR	SUNNY, MUDDY
04/06/2005	2	BLIND GRID	745	800	15	DAILY START, STOP	SET UP EQUIPMENT	GPS	NA	LINEAR	SUNNY, MUDDY

Note: Activities pertinent to this specific demonstration are indicated in highlighted text.

Date	No. of People	Area-Tested	Status Start Time	Status Stop Time	Duration min.	Operational Status	Operational Status - Comments	Track Method	Track Method=Other Explain	Pattern	Field Conditions
04/06/2005	2	BLIND GRID	800	935	95	COLLECTING DATA		GPS	NA	LINEAR	SUNNY, MUDDY
04/06/2005	2	BLIND GRID	935	940	5	DOWNTIME DUE TO EQUIP MAINT/CHECK	CHANGE BATTERY	GPS	NA	LINEAR	SUNNY, MUDDY
04/06/2005	2	BLIND GRID	940	1020	40	COLLECTING DATA		GPS	NA	LINEAR	SUNNY, MUDDY
04/06/2005	2	BLIND GRID	1020	1030	10	BREAK/LUNCH		GPS	NA	LINEAR	SUNNY, MUDDY
04/06/2005	2	BLIND GRID	1030	1150	80	COLLECTING DATA		GPS	NA	LINEAR	SUNNY, MUDDY
04/06/2005	2	BLIND GRID	1150	1300	70	BREAK/LUNCH		GPS	NA	LINEAR	SUNNY, MUDDY
04/06/2005	2	BLIND GRID	1425	1620	115	COLLECTING DATA		GPS	NA	LINEAR	SUNNY, MUDDY
04/06/2005	2	BLIND GRID	1620	1630	10	DAILY START, STOP	EQUIPMENT BREAKDOWN	GPS	NA	LINEAR	SUNNY, MUDDY
04/07/2005	2	BLIND GRID	745	800	15	DAILY START, STOP	SET UP EQUIPMENT	GPS	NA	LINEAR	CLOUDY, MUDDY
04/07/2005	2	BLIND GRID	800	1030	150	COLLECTING DATA		GPS	NA	LINEAR	CLOUDY, MUDDY
04/07/2005	2	BLIND GRID	1030	1035	5	BREAK/LUNCH		GPS	NA	LINEAR	CLOUDY, MUDDY
04/07/2005	2	BLIND GRID	1035	1200	85	COLLECTING DATA		GPS	NA	LINEAR	CLOUDY, MUDDY
04/07/2005	2	BLIND GRID	1200	1320	80	BREAK/LUNCH		GPS	NA	LINEAR	CLOUDY, MUDDY
04/07/2005	2	BLIND GRID	1320	1500	100	COLLECTING DATA		GPS	NA	LINEAR	CLOUDY, MUDDY
04/14/2005	2	BLIND GRID	1600	1630	30	DEMOBILIZATION		GPS	NA	LINEAR	CLOUDY, MUDDY

Note: Activities pertinent to this specific demonstration are indicated in highlighted text.

APPENDIX E. REFERENCES

1. Standardized UXO Technology Demonstration Site Handbook, DTC Project No. 8-CO-160-000-473, Report No. ATC-8349, March 2002.

2. Aberdeen Proving Ground Soil Survey Report, October 1998.

3. Data Summary, UXO Standardized Test Site: APG Soils Description, May 2002.

4. Yuma Proving Ground Soil Survey Report, May 2003.

APPENDIX F. ABBREVIATIONS

AEC	=	U.S. Army Environmental Center
APG	=	Aberdeen Proving Ground
ATC	=	U.S. Army Aberdeen Test Center
DC	=	direct current
ERDC	=	U.S. Army Corps of Engineers Engineering Research and Development Center
ESTCP	=	Environmental Security Technology Certification Program
EQT	=	Army Environmental Quality Technology Program
GPS	=	Global Positioning System
JPG	=	Jefferson Proving Ground
POC	=	point of contact
QA	=	quality assurance
QC	=	quality control
ROC	=	receiver-operating characteristic
SERDP	=	Strategic Environmental Research and Development Program
UXO	=	unexploded ordnance
YPG	=	U.S. Army Yuma Proving Ground

(Page F-2 Blank)

APPENDIX G. DISTRIBUTION LIST

DTC Project No. 8-CO-160-UXO-021

Addressee	No. of Copies
Commander U.S. Army Environmental Center ATTN: SFIM-AEC-ATT (Ms. Kimberly Watts)	1
(Ms. Bonnie Packer)	1
Aberdeen Proving Ground, MD 21010-5401	
ARM Group Inc. ATTN: (Mr. Jeffrey Leberfinger, P.G.) 1129 West Governor Road P.O. Box 797 Hershey, PA 17033	1
Commander U.S. Army Aberdeen Test Center ATTN: CSTE-DTC-SL-E (Mr. Larry Overbay)	1
(Library)	1
CSTE-DTC-AT-CS-R	1
Aberdeen Proving Ground, MD 21005-5059	
Defense Technical Information Center 8725 John J. Kingman Road, STE 0944 Fort Belvoir, VA 22060-6218	1

Secondary distribution is controlled by Commander, U.S. Army Environmental Center, ATTN: SFIM-AEC-ATT.

(Page G-2 Blank)