AD NO. ___________________________
DTC PROJECT NO. 8-CO-160-UXO-021
REPORT NO. ATC-8836

STANDARDIZED

UXO TECHNOLOGY DEMONSTRATION SITE

BLIND GRID SCORING RECORD NO. 213

SITE LOCATION:
U.S. ARMY YUMA PROVING GROUND

DEMONSTRATOR:
NAVAL RESEARCH LABORATORIES
CODE 6110 NAVAL RESEARCH LABORATORIES
WASHINGTON, DC 20375-5342

TECHNOLOGY TYPE/PLATFORM:
MULTI-SENSOR TOWED ARRAY DETECTION
SYSTEM (MTADS) (GEM-3)

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

JANUARY 2005

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

TABLE OF CONTENTS

PAGE

SECTION 4. TECHNICAL PERFORMANCE RESULTS

SECTION 5. ON-SITE LABOR COSTS

SECTION 6. COMPARISON OF RESULTS TO DATE

SECTION 7. APPENDIXES

SECTION 1. GENERAL INFORMATION

1.1 BACKGROUND

Technologies under development for the detection and discrimination of unexploded ordnance (UXO) 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 varies 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.

SECTION 2. DEMONSTRATION

2.1 DEMONSTRATOR INFORMATION

2.1.1 Demonstrator Point of Contact (POC) and Address

Address: Naval Research Laboratory
Code 6110
Naval Research Laboratory
Washington, DC 20375-5342

2.1.2 System Description (provided by demonstrator)

The Multi-Sensor Towed Array Detection System (MTADS) GEM-3 is composed of three 96-cm diameter frequency-domain electromagnetic interference (EMI) sensors mounted in a triangular array (fig. 1). The array is mounted on a 3.5-meter long platform that is pulled by the MTADS tow vehicle (fig. 1). The sensor-transmit electronics and signal analog to digitals (A/Ds) are located on the tow platform just in front of the sensor coils; the remaining sensor electronics are rack-mounted in the tow vehicle. Also mounted on the tow platform are three Global Positioning System (GPS) antennae and an International Measurement Unit (IMU).

Figure 1. Demonstrator's system, MTADS GEM-3.

Each of the three sensors in the array sequentially transmits a composite waveform made up of ten frequencies logarithmically spaced from 30-Hz to just over 20 kHz for one base period (1/30 s). Thus, only one complete cycle of the 30-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 counter-wound bucking coil. This sets up a "magnetic cavity" inside the bucking coil, in which a receive coil is placed. 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 (parts per million (PPM)) as both an in-phase and a quadrature component.

These 20 measured responses (in-phase and quadrature at ten frequencies) make up the EMI Spectrum of the buried targets. These spectra can be analyzed by fitting to empirical functions, comparing against known library spectra, or fitting to target response coefficients. All three of these analysis methodologies will be applied to the data collected in this demonstration, and their results will be compared.

2.1.3 Data Processing Description (provided by demonstrator)

The MTADS GEM-3 consists of three, 96-cm diameter sensors arranged in a triangle. The array is pulled by the MTADS tow vehicle over the site at approximately 3 miles per hour. Lane spacing is the width of the MTADS tow vehicle, approximately 1.75 meters. Data are recorded from the array at approximately 9.7 Hz. This results in a down-track sampling interval of ~15 cm and a cross-track sampling interval of 50 cm. For the measurements at APG, data will be recorded while traversing the test field in two orthogonal directions (roughly north to south and east to west). As part of the analysis, the extra classification performance (if any) that results from these extra data will be determined.

Individual sensors in the array are located using a three-receiver, real-time kinematics (RTK) GPS system, as shown in Figure 1. From this set of receivers, the position of the master antenna is recorded at 20 Hz, and the vectors to the other two antennae are recorded at 10 Hz. All positions are recorded at full RTK precision, ~2-5 cm. In addition, the output of a full 6-axis IMU at 80 Hz is recorded to give complementary information on platform pitch and roll. All sensor readings are referenced to the GPS PostPostscriptum (1-PPS) output so that the precision of the GPS measurements can be utilized to full advantage.

The individual data streams into the data acquisition computer, running a custom variant of the WinGEM program called WinGEMArray, are each recorded in a separate file. These individual data files, which share a root name corresponding to the date and time the survey was initiated, include three sensor data files, four GPS files (one containing the National Maritime Electronics Association (NMEA) GGK sentences corresponding to the position of the master antenna and an automatic volume recognition (AVR) sentence giving one of the vectors to the secondary antennae, a second containing the second AVR sentence, a third containing the universal time coordinated (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 in American Standard Code for Information Interchange (ASCII) format, and the IMU file mirrors the packed binary output of the IMU.

All of these files are transferred to the data analysis system using ZIP-250 disks and are checked for data quality and leveled; the position information is then applied to the sensor files. The result is a sequence of positioned measurements of the measured response at ten frequencies; this latter file is referred to as raw data.

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 (app E, ref 1). 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)

To ensure adequate system performance, three items must be checked daily: individual sensor response, timing accuracy of sensor measurements, and reliability of GPS positions. Before beginning survey work each day, the performance of each of the three sensors in the array is measured (after a 5-min warm-up) by presenting a ferrite rod and a standard sphere as targets. These test targets are mounted on a short, wooden block placed directly on the sensor coils. The resulting frequency-dependent signals are checked against standard values.

System timing accuracy is checked by making a back-and-forth traverse over a linear target at the beginning and end of each 1-hour survey file. These targets can be either a steel wire stretched between stakes or a small-diameter (1/2-in.) copper pipe placed on the ground adjacent to the survey area. ATC on-site personnel will determine the best target.

The data acquisition system gives the vehicle operator a continuous reading of the quality of the GPS fix. The standard procedure is to take only data with a GPS fix quality of 3 (RTK fixed) or 2 (RTK float) and a precision dilution of precision (PDOP) of 4 or less. Before arriving at the site each day, standard GPS planning software is used to calculate the number of satellites that will be visible to the receivers and the PDOP achievable minute-by-minute throughout the day. This allows GPS planning during periods of poor satellite availability and keeps inadvertent data, which would have to be discarded, from being recorded. Another important feature provided by GPS planning is the ability to take into account areas of restricted sky view (such as the tree line at one edge of the APG site). Past experience has shown that a brief period usually occurs each day, about 20 to 30 minutes, when good fixes can be obtained in even the most difficult environments. With planning, the system can be poised by the tree line ready to take data when the appropriate satellite alignment occurs.

Overview of QA. At the end of each 1-hour survey session, all survey data are transferred to the field data analyst for preliminary data quality checks. This process involves plotting the actual survey path as logged in the GPS files (color-coded by GPS fix quality) to ensure that GPS data of sufficient quality were obtained during the survey. Following this, the individual sensor files are examined for completeness and consistency. At this stage, sensor malfunctions, drifts, etc., are flagged and reported to the field crew for correction. The final objective for the field analyst is to calculate a position for each sensor reading and apply it to the reading. The mapped data files are then ready for analysis either in the field or at a later time.

2.1.6 Additional Records

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

2.2 YPG SITE INFORMATION

2.2.1 Location

YPG is located adjacent to the Colorado River in the Sonoran Desert. The UXO Standardized Test Site is located south of Pole Line Road and east of the Countermine Testing and Training Range. The Open Field range, Calibration Grid, Blind Grid, Mogul area, and Desert Extreme area comprise the 350- by 500-meter general test site area. The open field site is the largest of the test sites and measures approximately 200 by 350 meters. To the east of the open field range are the calibration and blind test grids that measure 30 by 40 meters and 40 by 40 meters, respectively. South of the Open Field is the 135- by 80-meter Mogul area consisting of a sequence of man-made depressions. The Desert Extreme area is located southeast of the open field site and has dimensions of 50 by 100 meters. The Desert Extreme area, covered with desert-type vegetation, is used to test the performance of different sensor platforms in a more severe desert condition/environment.

2.2.2 Soil Type

Soil samples were collected at the YPG UXO Standardized Test Site by ERDEC to characterize the shallow subsurface (<3 m). Both surface grab samples and continuous soil borings were acquired. The soils were subjected to several laboratory analyse, including sieve/hydrometer, water content, magnetic susceptibility, dielectric permittivity, X-ray diffraction, and visual description.

Two soil complexes are present within the site, Riverbend-Carrizo and Cristobal-Gunsight. The Riverbend-Carrizo complex is composed of mixed-stream alluvium, whereas the Cristobal-Gunsight complex is derived from fan alluvium. The Cristobal-Gunsight complex covers the majority of the site. Most of the soil samples were classified as either a sandy loam or loamy sand, with most samples containing gravel-size particles. All samples had a measured water content of less than 7 percent, except for two that contained 11 percent moisture. The majority of soil samples had water content between 1 and 2 percent. Samples containing more than 3 percent were generally deeper than 1 meter.

An X-ray diffraction analysis on four soil samples indicated a basic mineralogy of quartz, calcite, mica, feldspar, magnetite, and some clay. The presence of magnetite imparted a moderate magnetic susceptibility, with volume susceptibilities generally greater than 100 by 10-5 SI.

For more details concerning the soil properties at the YPG 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 YPG is included in Table 2.

TABLE 2. TEST SITE AREAS

Area	Description
Calibration Grid	Contains the 15 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 0.16-hectare (0.39-acre) site. The center of each grid cell contains ordnance, clutter, or nothing.

SECTION 3. FIELD DATA

3.1 DATE OF FIELD ACTIVITIES: 12 and 13 November 2003

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	1.93
Blind Grid	3.17

3.3 TEST CONDITIONS

3.3.1 Weather Conditions

A YPG weather station located approximately 1-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 the precipitation data represent 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, 2003	Average Temperature, ^{o}F	Total Daily Precipitation, in.
12 November	N/A	N/A
13 November	68.9	0.00
14 November	62.9	0.00
19 November	72.1	0.00

3.3.2 Field Conditions

The field conditions remained dry throughout the demonstration, with the exception of 12 November, when testing was delayed due to rain conditions.

3.3.3 Soil Moisture

Three soil probes were placed at various locations within the site to capture soil moisture data: Calibration, Mogul, and Desert Extreme areas. Measurements were collected in percent moisture and were taken twice daily (morning and afternoon) from five different soil depths (0 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 breakdown. The initial setup of equipment took 3 hours on 12 November 2003. No time was spent on daily setup on 13 and 14 November 2003. Daily breakdown time was minimal; the vehicle was simply parked, so no time was logged for this activity.

3.4.2 Calibration

The Naval Research Laboratories (NRL) spent the morning of 12 November 2003 setting up their equipment and preparing to survey the Calibration Lanes. This time is captured in setup/mobilization (para 3.4.1). NRL surveyed the Calibration Lanes on the morning of 13 November 2003 in 1-hour and 56 minutes.

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 not discussed either.

3.4.3.1 Equipment/data checks, maintenance. Equipment/data checks and maintenance took a total of 34 minutes throughout the course of surveying the Calibration Lanes and Blind Grid.

3.4.3.2 Equipment failure or repair. Three problems were associated with NRL equipment in the Calibration and Blind Grid areas. The MTADS GEM-3 required soldering of a torn transmitter wire, which took 20 minutes to complete. NRL removed and replaced the number one transmitter as well, which took 15 minutes. The final required repair was to replace a spark plug wire, which took 19 minutes to complete.

3.4.3.3 Weather. Because conditions were too wet on 12 November 2003 (the first day of the scheduled demonstration), the survey was delayed. After the NRL crew unloaded their equipment and ran a systems check, they left the range for the day. The next day's weather was fine for starting the data collection.

3.4.4 Data Collection

NRL spent 1-hour and 42 minutes collecting data in the Blind Grid. This time excludes break/lunches and downtimes described in paragraph 3.4.3.

3.4.5 <u>Demobilization</u>

NRL then conducted a demonstration of the Open Field. Therefore, demobilization did not occur until 19 November 2003, when the crew spent 2 hours and 20 minutes breaking down and packing up their equipment.

3.5 PROCESSING TIME

NRL submitted the raw data from the demonstration activities on the last day of the demonstration, as required. The scoring submittal data was also provided within the required 30-day time frame.

3.6 DEMONSTRATOR'S FIELD SURVEYING METHOD

NRL began surveying in the southeast corner of both the Calibration and Blind Grids. Both surveys were conducted in an east/west direction. NRL started the survey of the open field area in the northeast corner and conducted the survey by running parallel to the boundary angle on the east side of the open field. This allowed them to move from the northeast toward the southwest at an angle, which was faster and a more economical use of their time and equipment.

3.7 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.

The only issue with NRL was the brief delay when their MTADS GEM-3 had equipment problems, which the crew was able to repair in a short period of time.

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.

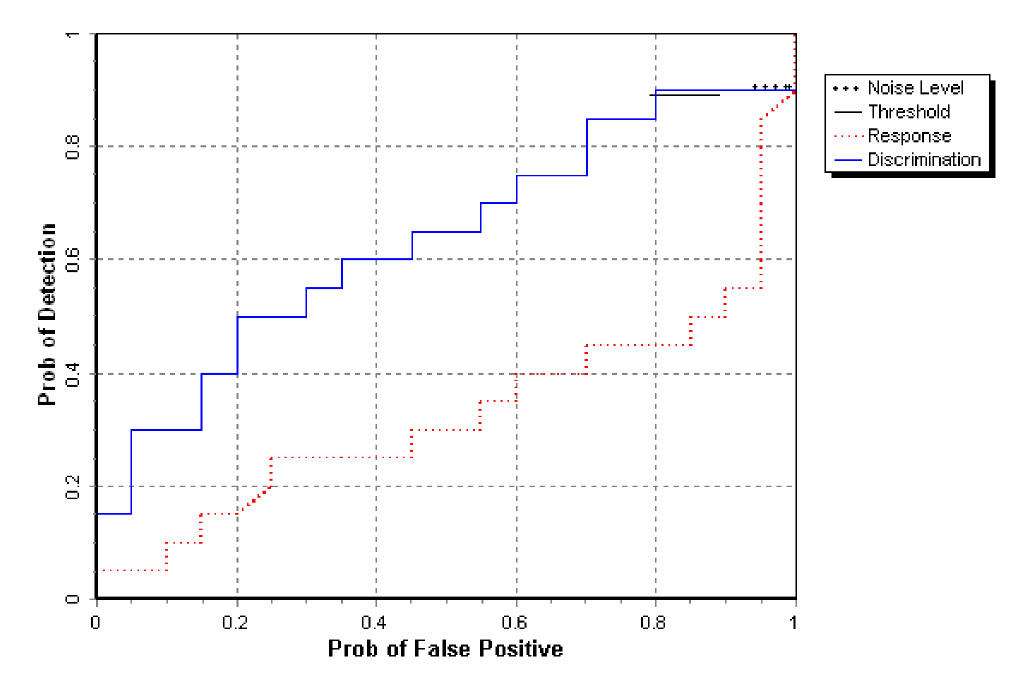

Figure 2. MTADS GEM-3 Blind Grid probability of detection for response and discrimination stages versus the respective probability of false positive over all ordnance categories combined.

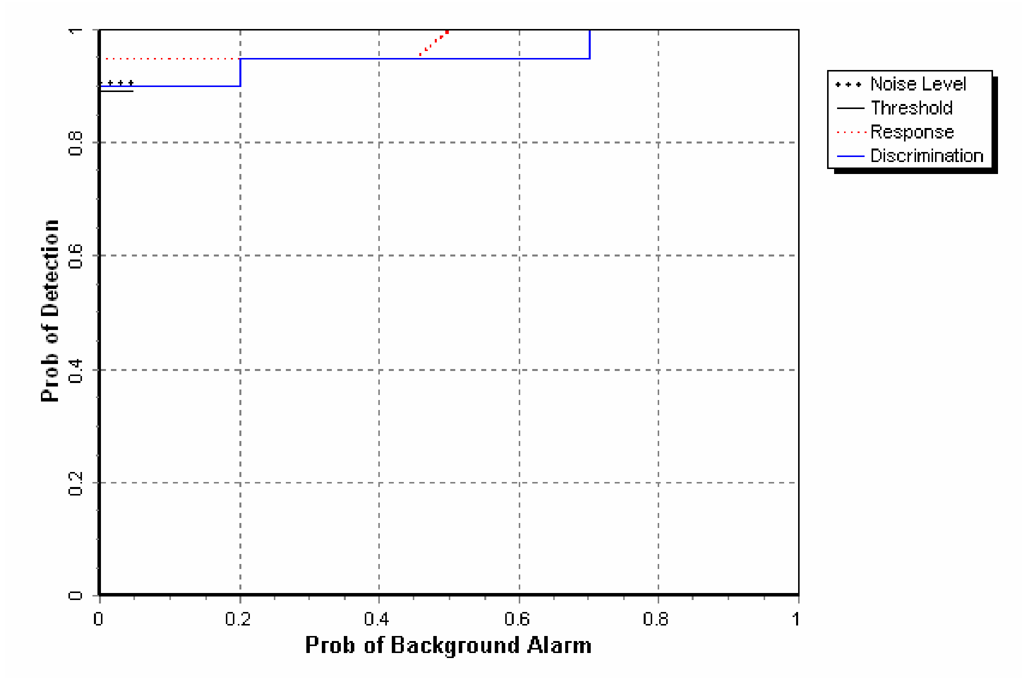

Figure 3. MTADS GEM-3 Blind Grid probability of detection for response and discrimination stages versus the respective probability of background alarm over all ordnance categories combined.

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.

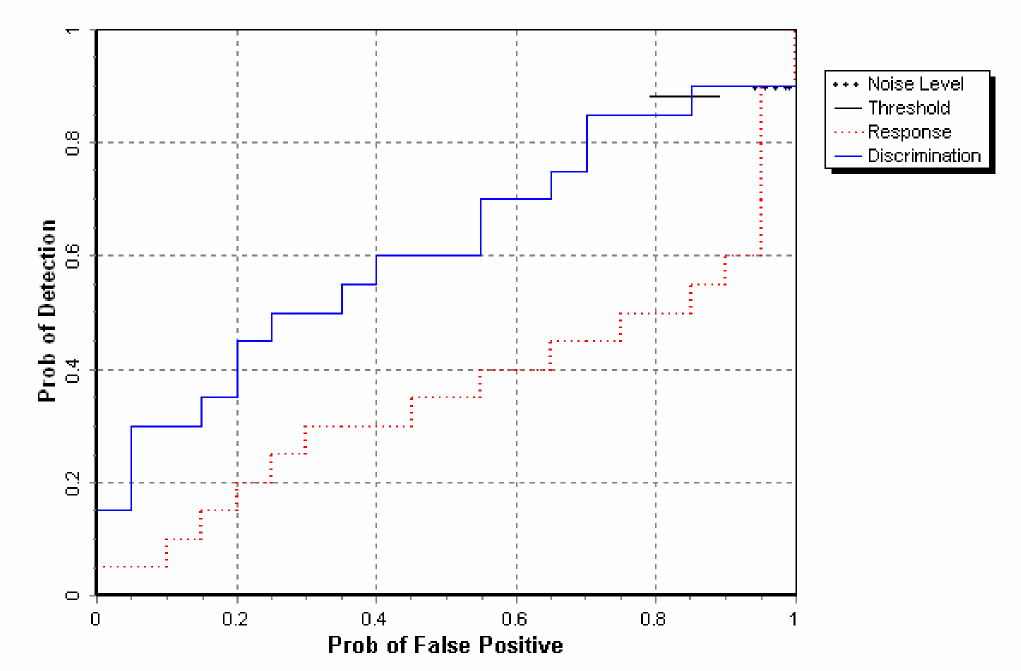

Figure 4. MTADS GEM-3 Blind Grid probability of detection for response and discrimination stages versus the respective probability of false positive for all ordnance larger than 20 mm.

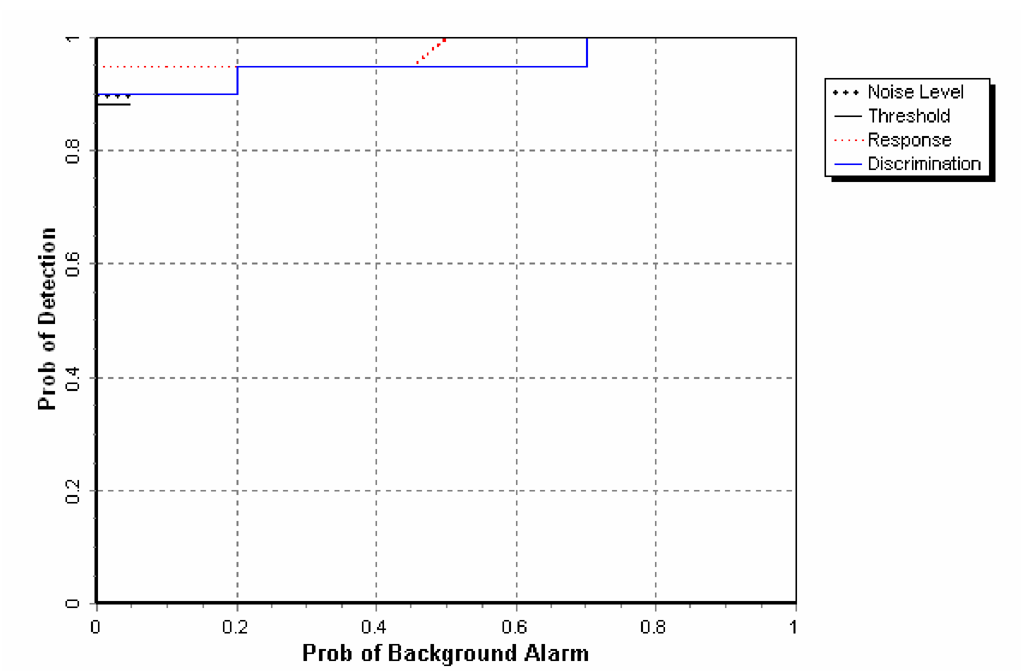

Figure 5. MTADS GEM-3 Blind Grid probability of detection for response and discrimination stages versus the respective probabilities of background alarm for all ordnance larger than 20 mm.

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 Appendix A for size definitions.) The results are relative to the number of ordnances emplaced. Depth is measured from the closest point of anomaly to the ground surface.

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 probability of false positive 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 MTADS GEM-3

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.

Notes: The response stage noise level and 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	0.99	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 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 three example items are: 20 mmP, 105 H, and 2.75 in, respectively.

TABLE 7. CORRECT TYPE CLASSIFICATION OF TARGETS CORRECTLY DISCRIMINATED AS UXO

Size	% Correct
Small	60.0
Medium	44.4
Large	15.4
Overall	47.0

Note: The demonstrator did not attempt to provide type classification.

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	-0.01	0.32

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	3.00	285.00
Data Analyst	1	57.00	3.00	171.00
Field Support	2	28.50	3.00	171.00
SubTotal				**$627.00**
CALIBRATION (NOT INCLUDING INITIAL SETUP)				
Supervisor	1	$95.00	1.93	183.35
Data Analyst	1	57.00	1.93	110.01
Field Support	2	28.50	1.93	110.01
SubTotal				**$403.37**
SITE SURVEY				
Supervisor	1	$95.00	3.17	301.15
Data Analyst	1	57.00	3.17	180.69
Field Support	2	28.50	3.17	180.69
SubTotal				**$662.53**

See notes at end of table.

TABLE 9 (CONT'D)

	No. People	Hourly Wage	Hours	Cost
DEMOBILIZATION				
Supervisor	1	$95.00	2.3	218.50
Data Analyst	1	57.00	2.3	131.10
Field Support	0	28.50	0.00	0.00
Subtotal				**$349.60**
Total				**$2,042.50**

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.

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.

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.

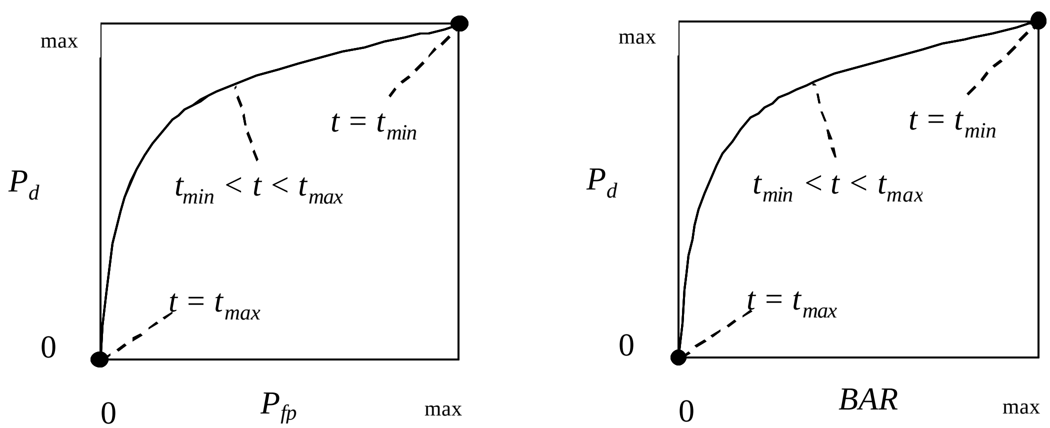

Figure A-1. ROC curves for open field-testing. Each curve applies to both the response and discrimination stages.

---

[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 t_{min}) 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 t_{min}). 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.

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 4).

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.

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 4).

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.

APPENDIX B. DAILY WEATHER LOGS

TABLE B-1. WEATHER LOG

Date	Time, EDST	Temperature, °F	RH, %	Precipitation, in.
13-Nov-03	15:00	68.5	44	0.00
13-Nov-03	16:00	68.9	44	0.00
13-Nov-03	17:00	68.5	44	0.00
13-Nov-03	18:00	66.0	49	0.00
13-Nov-03	19:00	63.3	61	0.00
13-Nov-03	20:00	61.9	67	0.00
13-Nov-03	21:00	59.5	75	0.00
13-Nov-03	22:00	59.7	73	0.00
13-Nov-03	23:00	57.9	77	0.00
13-Nov-03	24:00	55.9	85	0.00
14-Nov-03	1:00	55.6	82	0.00
14-Nov-03	2:00	54.3	84	0.00
14-Nov-03	3:00	52.3	88	0.00
14-Nov-03	4:00	51.3	90	0.00
14-Nov-03	5:00	50.4	90	0.00
14-Nov-03	6:00	50.0	91	0.00
14-Nov-03	7:00	49.8	89	0.00
14-Nov-03	8:00	50.3	91	0.00
14-Nov-03	9:00	55.9	79	0.00
14-Nov-03	10:00	57.9	77	0.00
14-Nov-03	11:00	61.9	68	0.00
14-Nov-03	12:00	65.5	56	0.00
14-Nov-03	13:00	68.7	46	0.00
14-Nov-03	14:00	70.2	41	0.00
14-Nov-03	15:00	71.1	39	0.00
14-Nov-03	16:00	71.4	38	0.00
14-Nov-03	17:00	70.7	35	0.00
14-Nov-03	18:00	67.8	40	0.00
14-Nov-03	19:00	64.8	52	0.00
14-Nov-03	20:00	61.3	68	0.00
14-Nov-03	21:00	60.6	65	0.00
14-Nov-03	22:00	58.8	71	0.00
14-Nov-03	23:00	56.1	83	0.00
14-Nov-03	24:00	55.6	82	0.00

APPENDIX C. SOIL MOISTURE

Daily Soil Moisture Logs

Demonstrator: NRL
Date: 13 November 2003
Times: 0710, 1730 hours

Probe Location:	Layer, in.	AM Reading, %	PM Reading, %
Calibration Area	0 to 6	1.9	1.9
	6 to 12	2.8	2.7
	12 to 24	3.7	3.7
	24 to 36	3.6	3.6
	36 to 48	4.0	4.0
Mogul Area	0 to 6	1.6	1.6
	6 to 12	2.7	2.6
	12 to 24	3.5	3.5
	24 to 36	4.0	4.0
	36 to 48	4.0	4.0
Desert Extreme Area	0 to 6	1.6	1.6
	6 to 12	2.5	2.4
	12 to 24	3.3	3.3
	24 to 36	3.9	3.9
	36 to 48	4.1	4.1

Date: 14 November 2003
Times: 0720, 1715 hours

Probe Location:	Layer, in.	AM Reading, %	PM Reading, %
Calibration Area	0 to 6	1.9	1.9
	6 to 12	2.6	2.6
	12 to 24	3.7	3.7
	24 to 36	3.6	3.6
	36 to 48	4.0	4.0
Mogul Area	0 to 6	1.6	1.6
	6 to 12	2.4	2.4
	12 to 24	3.5	3.5
	24 to 36	3.9	3.9
	36 to 48	4.0	4.0
Desert Extreme Area	0 to 6	1.8	1.8
	6 to 12	2.4	2.4
	12 to 24	3.3	3.3
	24 to 36	3.9	3.9
	36 to 48	4.1	4.1

Date: 17 November 2003
Times: 0655, 1715 hours

Probe Location:	Layer, in.	AM Reading, %	PM Reading, %
Calibration Area	0 to 6	1.8	1.8
	6 to 12	2.5	2.5
	12 to 24	3.7	3.7
	24 to 36	3.6	3.6
	36 to 48	4.0	4.0
Mogul Area	0 to 6	1.6	1.6
	6 to 12	2.3	2.3
	12 to 24	3.5	3.5
	24 to 36	3.9	3.9
	36 to 48	4.0	4.0
Desert Extreme Area	0 to 6	1.7	1.7
	6 to 12	2.3	2.3
	12 to 24	3.3	3.3
	24 to 36	3.9	3.9
	36 to 48	4.1	4.1

Date: 18 November 2003
Times: 0650, 1715 hours

Probe Location:	Layer, in.	AM Reading, %	PM Reading, %
Calibration Area	0 to 6	1.8	1.8
	6 to 12	2.5	2.5
	12 to 24	3.7	3.7
	24 to 36	3.6	3.6
	36 to 48	4.0	4.0
Mogul Area	0 to 6	1.6	1.6
	6 to 12	2.3	2.3
	12 to 24	3.5	3.5
	24 to 36	3.9	3.9
	36 to 48	3.9	3.9
Desert Extreme Area	0 to 6	1.6	1.6
	6 to 12	2.4	2.4
	12 to 24	3.3	3.3
	24 to 36	3.9	3.9
	36 to 48	4.1	4.1

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	
20031112	4	INTIAL SETUP	840	1110	150	SETUP/MOBILIZATION	SETTING UP EQUIPMENT FOR TESTING	NA	NA	NA	COOL	RAIN
20031113	4	INTIAL SETUP	710	740	30	SETUP/MOBILIZATION	SETTING UP EQUIPMENT FOR TESTING	NA	NA	NA	COOL	DRY
20031113	4	CALIBRATION LANE	740	815	35	COLLECTING DATA	SYSTEM WAS CALIBRATED FOR TESTING	GPS	NA	LINER	COOL	DRY
20031113	4	CALIBRATION LANE	815	857	42	COLLECTING DATA	RUNNING CALIBRATION LANE BIDIRECTIONAL EAST/WEST	GPS	NA	LINER	COOL	DRY
20031113	4	CALIBRATION LANE	857	917	20	DOWNTIME DUE TO EQUIP MAINT/CHECK	CHECKING/DOWNLOADING DATA	NA	NA	NA	COOL	DRY
20031113	4	CALIBRATION LANE	917	936	19	DOWNTIME DUE TO EQUIPMENT FAILURE	CHANGED SPARK PLUG WIRE	NA	NA	NA	COOL	DRY
20031113	4	BLIND TEST GRID	936	1040	64	COLLECTING DATA	RUNNING BTG BIDIRECTIONAL EAST/WEST	GPS	NA	LINER	COOL	DRY
20031113	4	BLIND TEST GRID	1040	1054	14	DOWNTIME DUE TO EQUIP MAINT/CHECK	CHECKING/DOWNLOADING DATA	NA	NA	NA	COOL	DRY
20031113	4	OPEN RANGE	1054	1202	68	COLLECTING DATA	RUNNING OPEN RANGE BIDIRECTIONAL EAST/WEST	GPS	NA	LINER	COOL	DRY
20031113	4	OPEN RANGE	1202	1330	88	BREAK/LUNCH	LUNCH	NA	NA	NA	WARM	DRY
20031113	4	BLIND TEST GRID	1330	1345	15	DOWNTIME DUE TO EQUIPMENT FAILURE	REMOVED AND REPLACED NUMBER ONE TRANSMITTER	NA	NA	NA	WARM	DRY
20031113	4	BLIND TEST GRID	1345	1425	40	DOWNTIME DUE TO EQUIPMENT FAILURE	SOLDERED TORN TRANSMITTER WIRES	NA	NA	NA	WARM	DRY
20031113	4	BLIND TEST GRID	1425	1452	27	COLLECTING DATA	SYSTEM WAS CALIBRATED FOR TESTING	GPS	NA	LINER	WARM	DRY
20031113	4	BLIND TEST GRID	1452	1503	11	COLLECTING DATA	RUNNING OPEN RANGE BIDIRECTIONAL EAST/WEST	GPS	NA	LINER	WARM	DRY
20031113	4	OPEN RANGE	1503	1521	18	DOWNTIME DUE TO EQUIP MAINT/CHECK	CHECKING/DOWNLOADING DATA	NA	NA	NA	WARM	DRY
20031113	4	OPEN RANGE	1521	1622	61	COLLECTING DATA	RUNNING OPEN RANGE BIDIRECTIONAL EAST/WEST	GPS	NA	LINER	COOL	DRY
20031113	4	OPEN RANGE	1622	1626	4	DOWNTIME DUE TO EQUIP MAINT/CHECK	CHECKING/DOWNLOADING DATA	NA	NA	NA	COOL	DRY
20031113	4	OPEN RANGE	1626	1721	55	COLLECTING DATA	RUNNING OPEN RANGE BIDIRECTIONAL EAST/WEST	GPS	NA	LINER	COOL	DRY

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	
20031113	4	OPEN RANGE	1721	1730	7	SETUP/MOBILIZATION	EQUIPMENT BREAKDOWN EOD	NA	NA	NA	COOL	DRY
20031114	4	OPEN RANGE	720	800	40	SETUP/MOBILIZATION	SETTING UP EQUIPMENT FOR TESTING	NA	NA	NA	COOL	DRY
20031114	4	OPEN RANGE	800	825	25	COLLECTING DATA	SYSTEM WAS CALIBRATED FOR TESTING	GPS	NA	LINER	COOL	DRY
20031114	4	OPEN RANGE	825	830	5	COLLECTING DATA	RUNNING SIGNATURE DATA ON M75	GPS	NA	LINER	COOL	DRY
20031114	4	OPEN RANGE	830	840	10	COLLECTING DATA	RUNNING SIGNATURE DATA ON 60 MM	GPS	NA	LINER	COOL	DRY
20031114	4	OPEN RANGE	840	850	10	SETUP/MOBILIZATION	SETTING UP EQUIPMENT FOR TESTING	NA	NA	NA	COOL	DRY
20031114	4	OPEN RANGE	850	930	40	COLLECTING DATA	RUNNING OPEN RANGE BIDIRECTIONAL EAST/WEST	GPS	NA	LINER	COOL	DRY
20031114	4	OPEN RANGE	930	936	6	DOWNTIME DUE TO EQUIP MAINT/CHECK	CHECKING/DOWNLOADING DATA	NA	NA	NA	COOL	DRY
20031114	4	OPEN RANGE	936	1045	69	COLLECTING DATA	RUNNING OPEN RANGE BIDIRECTIONAL EAST/WEST	GPS	NA	LINER	COOL	DRY
20031114	4	OPEN RANGE	1045	1050	5	DOWNTIME DUE TO EQUIP MAINT/CHECK	CHECKING/DOWNLOADING DATA	NA	NA	NA	COOL	DRY
20031114	4	OPEN RANGE	1050	1204	74	COLLECTING DATA	RUNNING OPEN RANGE BIDIRECTIONAL EAST/WEST	GPS	NA	LINER	WARM	DRY
20031114	4	OPEN RANGE	1204	1228	24	BREAK/LUNCH	LUNCH	NA	NA	NA	WARM	DRY
20031114	4	OPEN RANGE	1228	1236	8	COLLECTING DATA	SYSTEM WAS CALIBRATED FOR TESTING	GPS	NA	LINER	WARM	DRY
20031114	4	OPEN RANGE	1236	1320	44	COLLECTING DATA	RUNNING OPEN RANGE BIDIRECTIONAL EAST/WEST	GPS	NA	LINER	WARM	DRY
20031114	4	OPEN RANGE	1320	1331	11	DOWNTIME DUE TO EQUIPMENT FAILURE	LOST GPS, REPLACED GPS BATTERY	NA	NA	NA	WARM	DRY
20031114	4	OPEN RANGE	1331	1431	60	COLLECTING DATA	RUNNING OPEN RANGE BIDIRECTIONAL EAST/WEST	GPS	NA	LINER	WARM	DRY
20031114	4	OPEN RANGE	1431	1438	7	DOWNTIME DUE TO EQUIP MAINT/CHECK	CHECKING/DOWNLOADING DATA	NA	NA	NA	WARM	DRY
20031114	4	OPEN RANGE	1438	1600	82	COLLECTING DATA	RUNNING OPEN RANGE BIDIRECTIONAL EAST/WEST	GPS	NA	LINER	WARM	DRY
20031114	4	OPEN RANGE	1600	1605	5	DOWNTIME DUE TO EQUIP MAINT/CHECK	CHECKING/DOWNLOADING DATA	NA	NA	NA	WARM	DRY
20031114	4	OPEN RANGE	1605	1654	49	COLLECTING DATA	RUNNING OPEN RANGE BIDIRECTIONAL EAST/WEST	GPS	NA	LINER	WARM	DRY
20031114	4	OPEN RANGE	1654	1700	6	COLLECTING DATA	SYSTEM WAS CALIBRATED FOR TESTING	GPS	NA	LINER	WARM	DRY

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	
20031114	4	OPEN RANGE	1700	1715	15	SETUP/MOBILIZATION	EQUIPMENT BREAKDOWN EOD	NA	NA	NA	WARM	DRY
20031117	4	OPEN RANGE	655	735	40	SETUP/MOBILIZATION	SETTING UP EQUIPMENT FOR TESTING	NA	NA	NA	COOL	DRY
20031117	4	OPEN RANGE	735	805	30	COLLECTING DATA	SYSTEM WAS CALIBRATED FOR TESTING	GPS	NA	LINER	COOL	DRY
20031117	4	OPEN RANGE	805	908	63	COLLECTING DATA	RUNNING OPEN RANGE BIDIRECTIONAL EAST/WEST	GPS	NA	LINER	COOL	DRY
20031117	4	OPEN RANGE	908	911	3	DOWNTIME DUE TO EQUIP MAINT/CHECK	CHECKING/DOWNLOADING DATA	NA	NA	LINER	COOL	DRY
20031117	4	OPEN RANGE	911	1011	60	COLLECTING DATA	RUNNING OPEN RANGE BIDIRECTIONAL EAST/WEST	GPS	NA	LINER	COOL	DRY
20031117	4	OPEN RANGE	1011	1016	5	DOWNTIME DUE TO EQUIP MAINT/CHECK	CHECKING/DOWNLOADING DATA	NA	NA	NA	COOL	DRY
20031117	4	OPEN RANGE	1016	1123	67	COLLECTING DATA	RUNNING OPEN RANGE BIDIRECTIONAL EAST/WEST	GPS	NA	LINER	COOL	DRY
20031117	4	OPEN RANGE	1123	1127	4	DOWNTIME DUE TO EQUIP MAINT/CHECK	CHECKING/DOWNLOADING DATA	NA	NA	NA	COOL	DRY
20031117	4	OPEN RANGE	1127	1209	42	COLLECTING DATA	RUNNING OPEN RANGE BIDIRECTIONAL EAST/WEST	GPS	NA	LINER	WARM	DRY
20031117	4	OPEN RANGE	1209	1215	6	DOWNTIME DUE TO EQUIP MAINT/CHECK	CHECKING/DOWNLOADING DATA	NA	NA	NA	WARM	DRY
20031117	4	OPEN RANGE	1215	1413	118	COLLECTING DATA	RUNNING OPEN RANGE BIDIRECTIONAL EAST/WEST	GPS	NA	LINER	WARM	DRY
20031117	4	OPEN RANGE	1413	1423	10	DOWNTIME DUE TO EQUIP MAINT/CHECK	CHECKING/DOWNLOADING DATA	NA	NA	NA	WARM	DRY
20031117	4	OPEN RANGE	1423	1505	42	DOWNTIME DUE TO EQUIPMENT FAILURE	LOST GPS, REPLACED GPS BATTERY	NA	NA	NA	WARM	DRY
20031117	4	OPEN RANGE	1505	1602	57	COLLECTING DATA	RUNNING OPEN RANGE BIDIRECTIONAL EAST/WEST	GPS	NA	LINER	WARM	DRY
20031117	4	OPEN RANGE	1602	1605	3	DOWNTIME DUE TO EQUIP MAINT/CHECK	CHECKING/DOWNLOADING DATA	NA	NA	NA	WARM	DRY
20031117	4	OPEN RANGE	1605	1645	40	COLLECTING DATA	RUNNING OPEN RANGE BIDIRECTIONAL EAST/WEST	GPS	NA	LINER	WARM	DRY
20031117	4	OPEN RANGE	1645	1655	10	COLLECTING DATA	SYSTEM WAS CALIBRATED FOR TESTING	GPS	NA	LINER	WARM	DRY
20031117	4	OPEN RANGE	1655	1715	20	SETUP/MOBILIZATION	EQUIPMENT BREAKDOWN EOD	NA	NA	NA	WARM	DRY
20031118	4	OPEN RANGE	650	745	55	SETUP/MOBILIZATION	SETTING UP EQUIPMENT FOR TESTING	NA	NA	NA	COOL	DRY

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	
20031118	4	OPEN RANGE	745	820	35	COLLECTING DATA	SYSTEM WAS CALIBRATED FOR TESTING	GPS	NA	LINER	COOL	DRY
20031118	4	OPEN RANGE	820	821	1	DOWNTIME DUE TO EQUIP MAINT/CHECK	LOOSE TRANSMITTER CABLE	NA	NA	NA	COOL	DRY
20031118	4	OPEN RANGE	821	1100	159	DOWNTIME DUE TO EQUIPMENT FAILURE	STRIPPED, CUT, SOLDERED CABLE NUMBER 2 WIRES	NA	NA	NA	COOL	DRY
20031118	4	OPEN RANGE	1100	1105	5	DOWNTIME DUE TO EQUIPMENT FAILURE	R&R TRANSMITTER BOX TO TEST CABLE	NA	NA	NA	COOL	DRY
20031118	4	OPEN RANGE	1105	1110	5	DOWNTIME DUE TO EQUIPMENT FAILURE	REINSTALLED TRANSMITTER BOX NUMBER 2	NA	NA	NA	COOL	DRY
20031118	4	OPEN RANGE	1110	1148	38	COLLECTING DATA	SYSTEM WAS CALIBRATED FOR TESTING	GPS	NA	LINER	WARM	DRY
20031118	4	OPEN RANGE	1148	1250	62	COLLECTING DATA	RUNNING OPEN RANGE BIDIRECTIONAL EAST/WEST	GPS	NA	LINER	WARM	DRY
20031118	4	OPEN RANGE	1250	1255	5	DOWNTIME DUE TO EQUIP MAINT/CHECK	CHECKING/DOWNLOADING DATA	NA	NA	NA	WARM	DRY
20031118	4	OPEN RANGE	1255	1400	5	COLLECTING DATA	RUNNING OPEN RANGE BIDIRECTIONAL EAST/WEST	GPS	NA	LINER	WARM	DRY
20031118	4	OPEN RANGE	1400	1410	10	DOWNTIME DUE TO EQUIP MAINT/CHECK	CHECKING/DOWNLOADING DATA	NA	NA	NA	WARM	DRY
20031118	4	OPEN RANGE	1410	1537	87	COLLECTING DATA	RUNNING OPEN RANGE BIDIRECTIONAL NORTH/SOUTH	GPS	NA	LINER	WARM	DRY
20031118	4	OPEN RANGE	1537	1540	3	DOWNTIME DUE TO EQUIP MAINT/CHECK	CHECKING/DOWNLOADING DATA	NA	NA	NA	WARM	DRY
20031118	4	OPEN RANGE	1540	1640	60	COLLECTING DATA	RUNNING OPEN RANGE BIDIRECTIONAL NORTH/SOUTH	GPS	NA	LINER	WARM	DRY
20031118	4	OPEN RANGE	1640	1646	6	COLLECTING DATA	SYSTEM WAS CALIBRATED FOR TESTING	GPS	NA	LINER	WARM	DRY
20031118	4	OPEN RANGE	1646	1705	19	SETUP/MOBILIZATION	EQUIPMENT BREAKDOWN EOD	NA	NA	NA	WARM	DRY
20031119	4	OPEN RANGE	655	735	40	SETUP/MOBILIZATION	SETTING UP EQUIPMENT FOR TESTING	NA	NA	NA	COOL	DRY
20031119	4	OPEN RANGE	735	808	34	COLLECTING DATA	SYSTEM WAS CALIBRATED FOR TESTING	GPS	NA	LINER	COOL	DRY
20031119	4	OPEN RANGE	808	822	14	COLLECTING DATA	RUNNING OPEN RANGE BIDIRECTIONAL NORTH/SOUTH	GPS	NA	LINER	COOL	DRY

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	
20031119	4	OPEN RANGE	822	835	13	DOWNTIME DUE TO EQUIP MAINT/CHECK	SECURED GPS ANTENNA TO VEHICLE	NA	NA	NA	COOL	DRY
20031119	4	OPEN RANGE	835	855	20	COLLECTING DATA	RUNNING OPEN RANGE BI-DIRECTIONAL NORTH/SOUTH	GPS	NA	LINER	COOL	DRY
20031119	4	OPEN RANGE	855	859	4	DOWNTIME DUE TO EQUIP MAINT/CHECK	CHECKING/DOWNLOADING DATA	NA	NA	NA	COOL	DRY
20031119	4	OPEN RANGE	859	943	44	COLLECTING DATA	RUNNING OPEN RANGE BI-DIRECTIONAL NORTH/SOUTH	GPS	NA	LINER	COOL	DRY
20031119	4	OPEN RANGE	943	951	8	DOWNTIME DUE TO EQUIP MAINT/CHECK	CHECKING/DOWNLOADING DATA	NA	NA	NA	COOL	DRY
20031119	4	OPEN RANGE	951	1035	44	COLLECTING DATA	RUNNING OPEN RANGE BIDIRECTIONAL NORTH/ SOUTH	GPS	NA	LINER	COOL	DRY
20031119	4	OPEN RANGE	1035	1040	5	DOWNTIME DUE TO EQUIP MAINT/CHECK	CHECKING/DOWNLOADING DATA	NA	NA	NA	COOL	DRY
20031119	4	OPEN RANGE	1040	1055	15	COLLECTING DATA	RUNNING OPEN RANGE BIDIRECTIONAL NORTH/SOUTH	GPS	NA	LINER	COOL	DRY
20031119	4	OPEN RANGE	1055	1100	5	BREAK/LUNCH	BREAK	NA	NA	NA	COOL	DRY
20031119	4	OPEN RANGE	1100	1134	34	COLLECTING DATA	RUNNING OPEN RANGE BIDIRECTIONAL EAST/WEST	GPS	NA	LINER	WARM	DRY
20031119	4	OPEN RANGE	1134	1136	2	DOWNTIME DUE TO EQUIP MAINT/CHECK	CHECKING/DOWNLOADING DATA	NA	NA	NA	WARM	DRY
20031119	4	OPEN RANGE	1136	1155	19	COLLECTING DATA	RUNNING OPEN RANGE BI-DIRECTIONAL EAST/WEST	GPS	NA	LINER	WARM	DRY
20031119	4	OPEN RANGE	1155	1201	6	DOWNTIME DUE TO EQUIP MAINT/CHECK	CHECKING/DOWNLOADING DATA	NA	NA	NA	WARM	DRY
20031119	4	OPEN RANGE	1201	1232	31	COLLECTING DATA	RUNNING OPEN RANGE BIDIRECTIONAL NORTH/SOUTH	GPS	NA	LINER	WARM	DRY
20031119	4	OPEN RANGE	1232	1247	15	DOWNTIME DUE TO EQUIP MAINT/CHECK	CHECKING/DOWNLOADING DATA	NA	NA	NA	WARM	DRY
20031119	4	OPEN RANGE	1247	1330	43	COLLECTING DATA	RUNNING OPEN RANGE BIDIRECTIONAL NORTH/SOUTH	GPS	NA	LINER	WARM	DRY
20031119	4	OPEN RANGE	1330	1332	2	BREAK/LUNCH	BREAK	NA	NA	NA	WARM	DRY
20031119	4	OPEN RANGE	1332	1337	5	COLLECTING DATA	RUNNING OPEN RANGE BIDIRECTIONAL EAST/WEST	GPS	NA	LINER	WARM	DRY
20031119	4	OPEN RANGE	1337	1357	20	DOWNTIME DUE TO EQUIP MAINT/CHECK	CHECKING/DOWNLOADING DATA	NA	NA	NA	WARM	DRY
20031119	2	OPEN RANGE	1357	1615	138	DEMOBILIZATION	END OF TEST, TURNED IN DISK	NA	NA	NA	COOL	DRY

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. Data Summary, UXO standardized Test Site: YPG Soils Description, May 2003.

5. Practical Nonparametric Statistics, W.J. Conover, John Wiley & Sons, 1980, pages 144 through 151.

APPENDIX F. ABBREVIATIONS

A/D	=	analog to digital
AEC	=	U.S. Army Environmental Center
APG	=	Aberdeen Proving Ground
ASCII	=	American Standard Code for Information Interchange
ATC	=	U.S. Army Aberdeen Test Center
AVR	=	automatic volume recognition
BTG	=	Blind Test Grid
EMI	=	electromagnetic interface
EOD	=	explosive ordnance disposed
ERDC	=	U.S. Army Corps of Engineers Engineering Research and Development Center
ESTCP	=	Environmental Security Technology Certification Program
EQT	=	Army Environmental Quality Technology Program
HEAT	=	high-explosive, antitank
GPS	=	Global Positioning System
IMU	=	International Measurement Unit
JPG	=	Jefferson Proving Ground
MTADS	=	Multi-Sensor Towed Array Detection System
NMEA	=	National Maritime Electronics Association
NRL	=	Naval Research Laboratories
P_d	=	probability of detection
PDOP	=	precision dilution of precision
POC	=	point of contact
PPM	=	parts per million
PPS	=	PostPostscriptum
PVC	=	polyvinyl chloride
QC	=	quality control
PDA	=	personal digital assistant
ROC	=	receiver-operating characteristic
RTK	=	real time kinematics
SERDP	=	Strategic Environmental Research and Development Program
UTC	=	universal time coordinated
UXO	=	unexploded ordnance
YPG	=	U.S. Army Yuma Proving Ground