AD NO. ___________________________
DTC PROJECT NO. 8-CO-160-UXO-016
REPORT NO. ATC-9534

FINAL REPORT

FOR THE

SHALLOW WATER

UXO TECHNOLOGY DEMONSTRATION SITE

SITE LOCATION:
U.S. ARMY ABERDEEN PROVING GROUND

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

MAY 2008

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

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

DISTRIBUTION UNLIMITED, MAY 2008.

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
MAY 2008	Final	September 2005 - September 2007

4. TITLE AND SUBTITLE
FINAL REPORT FOR THE SHALLOW WATER UXO TECHNOLOGY DEMONSTRATION SITE

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Rowe, Gary

5d. PROJECT NUMBER
8-CO-160-UXO-016

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Commander
U.S. Army Aberdeen Test Center
ATTN: CSTE-DTC-AT-SL-E
Aberdeen Proving Ground, MD 21005-5059

8. PERFORMING ORGANIZATION REPORT NUMBER
ATC-9534

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Commander
U.S. Army Environmental Command
ATTN: IMAE-ATT
Aberdeen Proving Ground, MD 21010-5401

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

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
Same as item 8

12. DISTRIBUTION/AVAILABILITY STATEMENT
Distribution unlimited.

13. SUPPLEMENTARY NOTES
None.

14. ABSTRACT

The U.S. Army Environmental Quality Technology (EQT) Program issued a Broad Agency Announcement (BAA) in support of EQT User Requirement A (1.6.a) UXO Screening, Detection, and Discrimination, requesting a demonstration of technologies for detection and discrimination of unexploded ordnance (UXO). Eight systems, provided by six companies, demonstrated at the Aberdeen Test Center (ATC) Shallow Water UXO Technology Demonstration Site, Aberdeen Proving Ground (APG), Maryland. The goal of the BAA was to provide opportunities for demonstration not only for detection or discrimination technologies but also technologies that significantly improve sensing, positioning, analysis, and visualization. This report presents data for system-level performance.

15. SUBJECT TERMS
UXO Technology Demonstration Site, Shallow Water, MEC. EMI, EM, Geophex Ltd., Tetra Tech Inc., IT Jewell Inc., Concurrent Technologies Corporation (CTC), NAEVA Geophysical Inc., XTECH, 3Dgeophysics, AMEC Earth and Environmental

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

TABLE OF CONTENTS

SECTION 1. EXECUTIVE DIGEST

SECTION 2. SYSTEMS UNDER TEST

SECTION 3. SYSTEM MANEUVERABILITY EVALUATION

SECTION 4. SCORING PROCEDURES

SECTION 5. SYSTEM PERFORMANCE BY TEST AREA

SECTION 6. SITE REMEDIATION COMPARISON

SECTION 7. APPENDIXES

SECTION 1. EXECUTIVE DIGEST

1.1 SUMMARY

a. The shallow water site contains three main test (scoring) areas: blind grid, open water, and littoral zone. The blind grid represents a bounded survey area, i.e., determine what, if anything is at each of 644 specific coordinates. This area provides the best indication of a system's detection range (signal sensitivity and resolution) and discrimination ability (ability to differentiate between ordnance items and clutter). The open water and littoral zone are unbounded regions, i.e., a significantly greater number of possible coordinate locations to evaluate. These two areas provide an indication of the system's performance in an environment more representative of the real world. The set of known items (GT targets) in both the open water and littoral zone contains, as a subset, the complete set of targets in the blind grid. This provides a more meaningful indication of how the additional variables encountered in the unbounded areas impact system performance. Water depths in these areas range from 1 to 10 feet.

b. There are two other sections in the shallow water site, the calibration lane and the deeper water area. The calibration lane contains the same ordnance items used as the GT targets throughout the test site. Prior to testing, the demonstrator is given the identity of each calibration item along with its coordinates and burial depths. This information is typically used for instrument calibration purposes or for the development of a reference target library for this site to use in the discrimination process. The deeper water area provides limited data on the system's probability of detection in water depths ranging from 10 to 14 feet.

c. The data in this report represent system-level performance. The structure of this evaluation was not designed to answer component-level questions. Each demonstrator operated his system and conducted the survey in accordance with the individual test plans submitted during the Broad Agency Announcement[1] (BAA) selection process. The final dig list produced by each demonstrator is the product of different components applied in different ways, variations in instrumentation calibration techniques, quality assurance and control methods, and data collection strategies. This does not include the variations in the data reduction and analysis methods.

d. Eight systems, provided by six companies, demonstrated under the BAA. During this program, each system was evaluated first as a metal detector, i.e., how well it could find metal objects and then on the ability of the associated discrimination process to segregate ordnance from clutter. Not all systems were able to survey all test areas.

e. This report shows that there are three areas in the design of shallow water systems that can be improved to increase the probability of detecting and identifying ordnance items:

(1) The maneuverability of the systems must allow the sensors to get within detection range of the areas of interest (section 3).

(2) The sensors need to be located at a distance from the surface of interest that provides sufficient signal resolution for both the detection and discrimination processes (para 5.2.5).

---

[1] . Broad Agency Announcement W91ZLK-04-R-0001, Unexploded Ordnance (UXO) Detection and Discrimination Demonstrations, 18 November 2004

(3) The Global Positioning System (GPS) must be able to locate the GT items with greater accuracy (paras 5.3.5.2 and 5.4.5.2).

f. The summary analyses of the detection and discrimination results by surveyor, across all test areas, are in section 5.6.

g. A standardized procedure was used to compare survey costs between demonstrators. The on-site costs to survey this 6.9-acre site ranged from $5.8 to $12.4K. Additional details are in section 6.1.

h. Numerous software packages and analytical approaches were used in the discrimination process with limited degrees of success. Improvements in the discrimination process have the greatest potential for both cost savings and risk reduction. The virtual site remediation analysis in section 6.2 provides an overview of the difference in site remediation before and after the discrimination process was applied to the metal detection (response) data.

1.2 BACKGROUND

a. Shallow water unexploded ordnance (UXO) detection and discrimination do not have the long history that ground-based detections have. Land-based UXO remediation has been ongoing for years with multiple companies using multiple sensors/systems and obtaining a variety of results. The ground impact areas have been classified by soil types and other geophysical features. Excavations at these sites have helped characterize the distribution, penetration depth, and orientation of the UXO hidden below the surface. In contrast, shallow water remediation has a short history, involving a limited number of companies and systems, limited underwater clearance attempts, no standard site classification scheme, and limited information on UXO distribution in a marine environment. The first attempt to detect ordnance in a shallow water test site is recorded in the Mare Island report[2].

b. The design of the ATC Shallow Water UXO Technology Demonstration Site incorporates the lessons learned from both the ground- and water-based histories, educated assumptions, engineering judgment, and established test methodologies. The objective was to construct a marine environment that provides a limited number of challenges for the survey platforms and a statistically valid means of measuring their detection and discrimination capabilities. Testing at this site is independently administered and analyzed by the government.

1.3 TESTING AUTHORITY

The Standardized UXO Technology Demonstration Site Program is a multiagency program headed by the U.S. Army Environmental Command (USAEC). ATC and the U.S. Army Corps of Engineers, Engineer Research and Development Center (ERDC) provide programmatic support. The Environmental Security Technology Certification Program (ESTCP), the Strategic Environmental Research and Development Program (SERDP), and the Army's Environmental Quality Technology (EQT) Program provided funding and support for this program.

---

[2] Validation of Detection Systems (VDS) Test Program Report, Mare Island Naval Shipyard (MINS), Vallejo, California. Prepared by the Environmental Chemical Corporation under Navy Contract No. N62742-98-D-1809, CTO 0001, 7 July 2000.

1.4 OBJECTIVES

The objectives of the testing conducted at the Shallow Water Standardized UXO Technology Demonstration Site are to evaluate the detection and discrimination capabilities of existing and emerging technologies and systems in a shallow water environment. Specifically:

a. To determine the demonstrator's ability to survey a shallow water area, analyze the survey data, and provide a prioritized target list in a timely manner.

b. To determine both the detection and discrimination effectiveness under realistic scenarios that vary ordnance, clutter, and bathymetric conditions.

c. To determine cost, time, and labor requirements needed to operate the technology.

1.5 ABERDEEN PROVING GROUND SHALLOW WATER SITE INFORMATION

a. Location. The Aberdeen Area of APG is located in the northeast portion of Maryland on the western shore of the Chesapeake Bay in Harford County. The Shallow Water Test Site is located within a controlled range area of APG.

b. Soil Type. The area chosen for the shallow water test site was known as cell No. 3 in a dredge-spoil field. The cell bottom is composed primarily of sediment removed from the nearby Bush River. This is a freshwater site.

c. Test Areas.

(1) The test site contains five areas: calibration grid, blind test grid, littoral, open water, and deeper water. A schematic of the test site showing the various test areas is shown in Figure 1-1. A detailed description of each area is in section 5.

(2) The water depth at this facility during testing is maintained such that the calibration and blind grid areas are under 8 feet (2.4 m) of water. The test site is approximately 2.8 hectares (6.9 acres) in size at this water level.

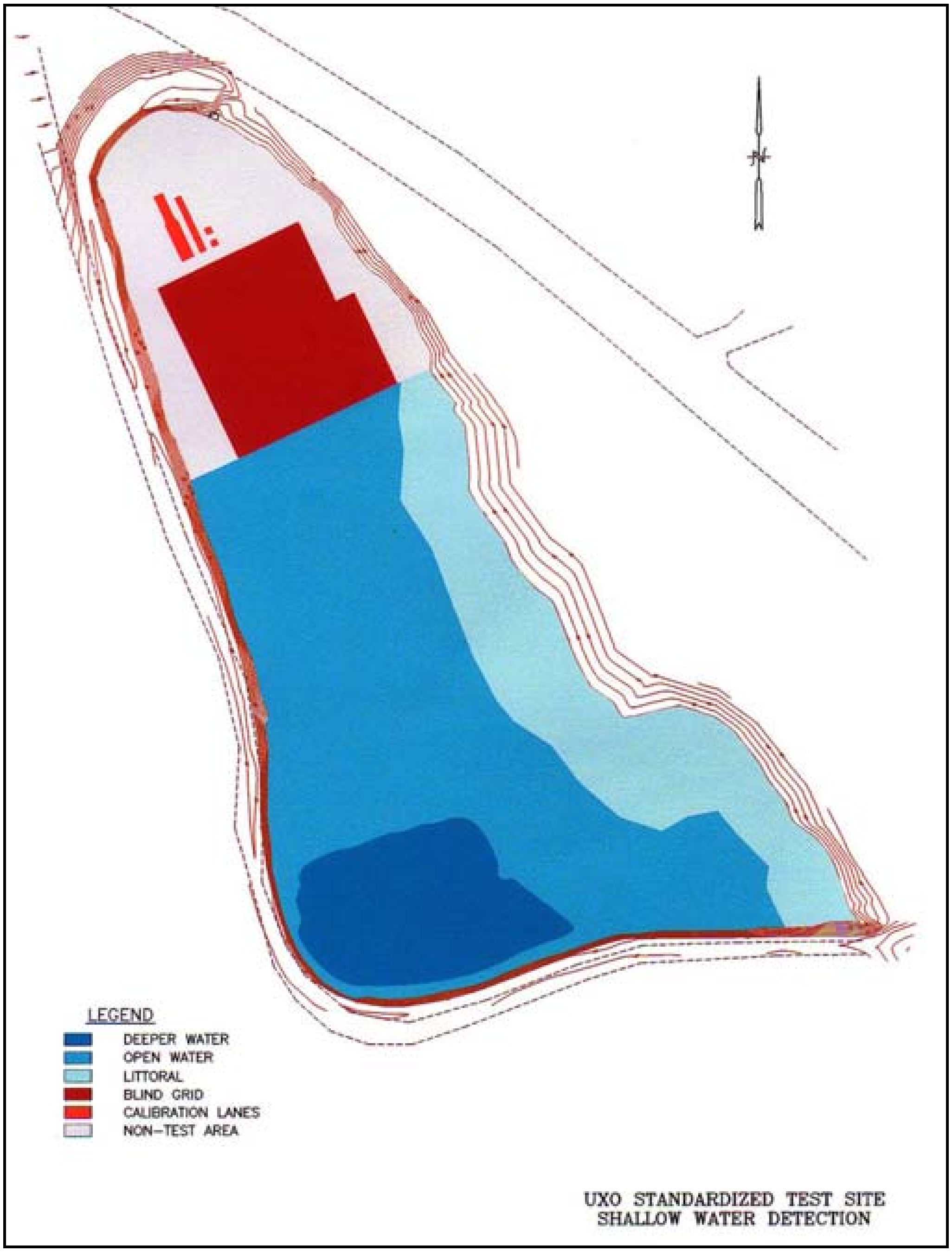


Figure 1-1. Shallow Water Test Site areas.

1.6 GROUND TRUTH TARGETS

a. The GT is composed of both inert ordnance and clutter items. The inert ordnance items are listed in Table 1-1. None of these items had been fired or degaussed.

TABLE 1-1. INERT ORDNANCE TARGETS

Description	Length, mm	Diameter, mm	Aspect Ratio, W/L	Weight, grams
40-mm L70 projectile	208	40	0.1923	965
60-mm mortar M49A2	185	60	0.3243	975
81-mm mortar M374	528	81	0.1534	3,969
81-mm mortar M821	510	81	0.1588	3,338
105-mm projectile M1	445	105	0.2360	13,834
155-mm M107 projectile	684	155	0.2266	41,731
8-in. M104/106	856	203	0.2371	89,811
Note: The non-availability of a single model of 81-mm mortars in sufficient quantities for test purposes necessitated grouping two models into one ordnance category. L = Length (mm) W = Width (mm)				

b. Clutter items fit into one of three categories: ferrous, nonferrous, and mixed metals. The ferrous and nonferrous items are further divided into the three weight zones shown in Table 1-2 and are distributed throughout all test areas. Most of this clutter is composed of ordnance components; however, industrial scrap metal and cultural items are present as well. The mixed-metals clutter is composed of scrap ordnance items or fragments that have both a ferrous and nonferrous component and could reasonably be encountered in a range area. The mixed-metals clutter is in the open water area only.

TABLE 1-2. CLUTTER WEIGHT RANGES

Clutter Type	Weight Range, grams		
	Small	Medium	Large
Ferrous	10 to 510	511 to 2,200	> 2,201
Nonferrous	10 to 270	275 to 800	> 801
Mixed metals	Weights ranged between 680 and 49,000 grams.		

SECTION 2. SYSTEMS UNDER TEST

Each demonstrator submitted a detailed test plan as part of the review process that led to accepting each system into this evaluation process. The descriptions of the various systems were condensed from those test plans.

2.1 GEOPHEX, LTD.

2.1.1 Detection Instrumentation

The GEM-3 array sensor is a continuous-wave frequency-domain electromagnetic induction (EMI) sensor that uses a hybrid current waveform to provide simultaneous multifrequency (typically ten log spaced) energy in the 90-Hz to 90-kHz band, with each receiver (Rx) digital signal processor performing digital Fourier transforms at the selected frequencies. A reference coil provides primary field reference (amplitude and phase) for Rx output normalization of the primary field-generated electromagnetic force. The sensors have a primary-field nulling scheme using a secondary concentric transmitter coil in series with the primary transmitter coil but current flowing in the opposite sense. Each sensor fires in sequence so that they do not interfere with each other.

2.1.2 Survey Platform

a. Two different platforms were required to survey the majority of the test site. The main platform has the sensor array mounted on a sled that has cement-filled polyvinyl chloride (PVC) runners for bottom-riding operation. The sled attaches to a solid-hull powerboat using a 40-foot fiberglass shaft (fig. 2-1). This configuration is used where the water is greater than a few feet deep and the boat/sensor combination has room to maneuver.

Figure 2-1. Geophex's towed sled and the vessel used to pull it.

b. A GPS-based navigation system determined the sensor position. The tow shaft extends up past the boat mount to provide mounting locations for dual Differential Global Positioning System (DGPS) antennas, several feet apart. Dual DGPS points allow for back-extrapolation, which provides both the pole angle from horizontal as well as heading.

c. A second platform was required in areas where the water was too shallow for the solid-hulled boat or space was too limited for the towed system to maneuver (fig. 2-2). This second platform used the same sensor array, mounted in a wooden frame and suspended approximately 3 feet below a small inflatable pontoon boat. Positioning again relies on dual DGPS antennas.

Figure 2-2. Geophex's second shallow water platform.

d. A third sensor configuration that was available at the test site consisted of an identical GEM-3 array housed in positive-buoyancy fiberglass housing. The concept was that of a floating array, towed behind the pontoon boat, along the shoreline or in very shallow water. This configuration was not evaluated.

2.1.3 Data Processing Description

a. The target detection process combines multifrequency data into a single detection channel designed to respond particularly to metal targets and not to geologic anomalies. Several were used including the sum of all quadrature channels, the difference between high frequency and low frequency in-phase channels, the sum of the absolute differences of quadrature channels between all frequency pairs and the inverse log (frequency) weighted total apparent conductivity. The selected detection channel forms the response stage. The DGPS georeferenced detection channel data are processed with an automatic anomaly picker that identifies target anomalies above a specified threshold, excluding single-point anomalies and overlapping secondary anomalies.

b. Georeferencing uses the time stamps to interpolate 15- or 30-Hz GEM-5 positions between 1-Hz DGPS fixes and the position for each sensor from spatial interpolation between two DGPS antennas. The raw DGPS latitude/longitude fixes are transformed to Universal Transverse Mercator (UTM) during the post-processing.

c. Target identification and classification (clutter discrimination) use a normalized matching of the multifrequency spectra to a library of known UXO spectral responses. The matching scheme fits an unknown target to the best-fit linear combination of the longitudinal (sensor axis along target long axis) and transverse (sensor axis perpendicular to target long axis) response spectra, allowing for a frequency-independent background in-phase response for magnetic soils. The goodness-of-fit to the best-fitting item is mapped into a confidence ranking from 0 (definite clutter) to 10 (definite UXO) with 5 corresponding to the clutter misfit threshold. The confidence ranking forms the discrimination stage.

2.1.4 ATC Survey Comments

The main survey platform is a slightly modified version of the system built in 2003 to survey an underwater area adjacent to Mare Island located within the San Francisco Bay. As anticipated, this configuration was too large to work well along the contours of the bank and in the shallow areas of the pond. Moving all the instrumentation to the second platform increased the area surveyed. Both platforms experienced minor structural or instrument interference problems. All problems were resolved during the survey.

2.2 TETRA TECH EC, INC.

2.2.1 Detection Instrumentation

This system uses both acoustic and magnetic geophysical technology. The acoustic technologies consist of a Sound Metrics Corporation dual-frequency imaging sonar and Specialty Devices, Inc., multifrequency subbottom profiler. The two acoustic systems identify metal objects on the pond bottom as well as in the shallow pond sediments (through acoustic reflectivity). The magnetic technology consists of GEM system's optically pumped potassium gradiometers (GSMP-40). The magnetic gradiometer system (MGS) geometry consists of three sensors in a triangular configuration and one sensor trailing the triangular array in a separate horizontal plane. This geometric design of the array allows measuring the total magnetic field for each sensor as well as six magnetic gradients. Four analytic signal measurements in two geometric planes are automatically calculated from the total field and gradient measurements. The use of the four analytic signal combinations delineates complex magnetic anomalies (e.g., representative of cluttered areas) into their individual constituents so that the anomaly locations are more representative of the individual items present. The magnetic technology confirms the visual indications from the acoustic systems and provides information on the presence of ferrous objects that are potentially out of the detection window of the acoustic systems (i.e., buried in the pond sediments).

2.2.2 Survey Platform

The marine survey platform consists of two 12-foot fiberglass Jon boats attached together, side by side, with fiberglass supports (fig. 2-3). There is a 2-foot space between the boats where the two acoustic technologies are mounted. GPS antennas mount on fiberglass poles that extend to a height of approximately 4 to 5 feet. The MGS array is on a third fiberglass vessel that is pushed 15 to 20 feet in front of the joined Jon boats. This platform is propelled by dual electric motors, one mounted on each Jon boat.

Figure 2-3. Tetra Tech shallow water UXO detection platform.

2.2.3 Data Processing Description

a. The high frequency imaging sonar (HFIS) data visually identifies the characteristics (length, orientation, overall visual properties, and coordinate position) of items on the pond bottom. The multifrequency subbottom profile (MFSBP) sonar data identifies the coordinates, distance, and acoustic reflectivity of metal items that exist above and below the pond sediments. The MGS data provide information on the characteristics of the ferrous metal items present in terms of their distance, ferrous mass, and magnetic dipole direction.

b. The processed data are evaluated with respect to the anomaly characteristics measured by each sensor used (e.g., signal intensity, visual identification through HFIS, anomaly size and shape, signal gradients, noise, and spatial sample density) to identify UXO-like items.

c. Geosoft Oasis Montaj® is used as the data interpretation platform. Color-coded images of the MGS and MFSBP data are generated and then compared with the coordinate locations of the items visually identified using the HFIS data. The coordinate location of each item is digitized and then classified by the interpreter to generate a dig sheet in the required format.

2.2.4 ATC Survey Comments

a. Overall, the design of the three-boat system and the instrumentation distribution among the boats were well thought out. This design appears best suited for surveying shallow or littoral regions.

b. The two electric motors that propelled the three-boat system did not have enough thrust to keep the boat on track in windy conditions. Replacing the two electric motors with a single 3.3-horsepower gasoline engine solved this problem.

c. Data files from the imaging sonar along with the minimally processed (raw) magnetometer data were provided at the completion of the survey on recordable compact discs (CD-R). The final data submission (dig list) was based on magnetometer data only. Tetra Tech provided the following explanation:

d. “We were unable to fully use all of the data from the five frequency sub bottom profiler because of software limitations for that particular system (we requested data be available in a certain SEG-Y format for all five frequencies - the vendor is still trying to correctly get this task completed by writing new software using a third party vendor). During data analysis, we used the system depth information and the current software ‘depthpic’ … to view the different frequencies, determine the approximate pond bottom, and correlate this information with the magnetic data. When (and if) the software can convert the multi-frequency data correctly, I expect that we have sufficient x-y data coverage of the area and expanded software analysis tools (as compared to depthpic software) to increase the overall usability of the data.

e. We did use the data from the DIDSON imaging sonar to select potential items and correlate with the magnetic data, however, it appears that there are extremely few or no metal items that were intended to exist on the pond bottom; however, we do not know this for sure since we do not have the truth data. Basically, the imaging sonar is (and was) implemented to detect items on the pond bottom[3] .”

f. Tetra Tech’s pairing of acoustic and metal-detecting instrumentation is a unique approach to underwater munitions and explosives of concern (MEC) detection. Problems with acoustic data interpretation, as described above, and detection limitations of the magnetometer array, as described later in this report, show that this particular application of a dual-sensor system was not very successful in terms of underwater MEC detection.

---

[3] . Email, SWDS Scoring, 6 July 2006, sent from Mr. Timothy Deigan (timothy.deigan@tteci.com) to Mr. Gary Rowe (gary.rowe@atc.army.mil).

2.3 IT JEWELL, INC.

2.3.1 Detection Instrumentation

IT Jewell surveyed the shallow water site twice, once using a Geonics 882 (G882) cesium-based marine magnetometer and a second time using the MM Explorer Mini proton-based magnetometer. Quality assurance (QA) and quality control measures ensured that factors such as site conditions, operator experience, survey methods, navigation, survey platform, data processing equipment, and components were consistent between the two surveys. The structure of this demonstration allowed not only for the survey of the test site but also for a side-by-side comparison of the two magnetometers.

2.3.2 Survey Platform

a. The survey platform consisted of a rectangular floating frame constructed from PVC pipe (fig. 2-4). Near the center of the frame is a cross member that holds the magnetometer approximately 0.5 meter below the water's surface and supports the GPS antenna approximately 0.5 meter above the surface. This platform is pushed through the survey area using a plastic Pond Prowler® pontoon boat powered by an electric motor.

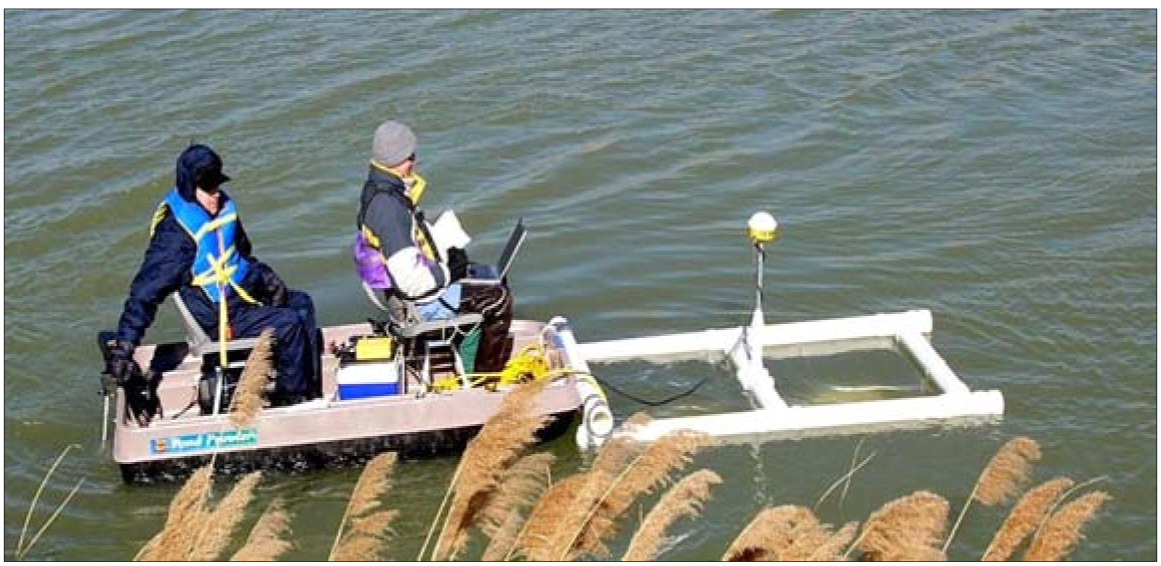


Figure 2-4. IT Jewell's shallow water survey platform.

2.3.3 Data Processing Description

a. Response Stage. Data collected where the boat turned around were typically removed from the analytical process. Total magnetic intensity was reviewed for noise, using a 4th horizontal derivative developed as an indication of noise content in magnetic data. Suitable convolution filters smoothed the data. A 5-point Hanning window was used on the G882 (0.5 sec), and a 3-point Hanning window (3 sec) was used on the MM Explorer Mini data. The data were manually corrected for diurnal variation by fitting a first-order polynomial to the data to remove any first-order tilting. Finally, the data were plotted and reviewed for each of the two sensor types.

b. Discrimination Stage. Filtering the magnetic data using a high-pass 1D Fast Fourier Transform (FFT) filter algorithm removed geologic signals. The design of the filter also enhances the signal over the calibration range for application to the rest of the data. Signals with wavelengths longer than 10 meters are rejected; wavelengths based on items picked using the calibration points provided by the Army are retained. The analytic signal, calculated from the processed data as the sum of the gradients, moves the peak analytic signal over the magnetic source. Anomalies are manually picked from individual profiles while reviewing the total intensity, high pass data, and analytic signal. Targets below the noise thresholds derived from the above calculations are rejected. Clutter was also largely included in this. Target positions, total magnetic intensity, and analytic signal (discrimination value) were tabled. Targets were finally discriminated, based on the range of observed data over the calibration range for each of the two sensor types.

2.3.4 ATC Survey Comments

a. This very basic one-sensor system was easy and inexpensive to transport, set up, and operate. Positioning the magnetometers approximately 0.5 meter below the surface and using a shallow-draft boat allowed this system to maneuver and survey in most of the littoral zone. However, the depth of the sensor became a detriment to the system in the deeper water areas (para. 5.2.5).

b. In the interest of evaluating as many shallow water detection systems as possible, and realizing the opportunity to have the data from two different sensors collected by and processed using one source, this approach had the potential to yield meaningful site- and sensor-specific data. Based on the site survey plan, many of the test variables needed for a valid side-by-side comparison appeared controlled to an acceptable level.

2.4 CONCURRENT TECHNOLOGIES CORPORATION (CTC)

2.4.1 Detection Instrumentation

CTC fielded a four-sensor array based on Foerster FEREX® 4.032 fluxgate vertical gradient magnetic sensor technology. The Foerster system is an off-the-shelf commercially available item that includes both the hardware necessary to collect and store the survey data and the software to interpret the collected information.

2.4.2 Survey Platform

Four sensors are placed in PVC pipe; the pipes were mounted in a wooden framework. This framework held the sensors along a single plane, separated by a horizontal distance of 0.5 meter. This array attached to the bow of an aluminum Jon boat using wooden brackets (fig. 2-5). The design of the wooden components allows the sensor array to be raised or lowered based on the water depth in which the system is operating. An electric outboard motor powers the 16-foot Jon boat. Sensor positioning relies on Trimble 5700 DGPS technology.

Figure 2-5. CTC shallow water survey platform in the raised and lowered positions.

2.4.3 Data Processing Description

a. During the data acquisition process, a FEREX® data logger records the signals from the four magnetometers and the DGPS position data. Foerster DATALINE® 4.800 software converts the magnetometer data to units of nanoTesla and then associates the signal strength with a DGPS location.

b. Foerster was responsible for the magnetometer data reduction and analysis. The DATALINE® 4.800 software provides a quality factor (0 - 100) to characterize the performance of the dipole-fit routine for each object calculation. The quality factor is associated with a volume/diameter calculation and a visual evaluation of the magnetic anomaly map. Using both numerical values produced by the software and a visual interpretation of the dipole on the anomaly map, the analyst determines if an object is clutter or an item of interest (ordnance).

2.4.4 ATC Survey Comments

a. This is the only boat-mounted system tested with the ability to vary the depth of the sensors with the water depth. In theory, keeping the magnetometers a uniform depth from the bottom should provide a more consistent signal response, leading to better detection and discrimination results.

b. Having a variable sensor depth also increases the maneuverability and capability of the system as the water levels change.

2.5 NAEVA GEOPHYSICAL, INC. (XTECH)

2.5.1 Detection Instrumentation

A multisensor underwater system using two Geonics EM61 MK2 (underwater coils) electromagnetic metal detectors.

2.5.2 Survey Platform

The deployed system consists of two underwater coils mounted side by side on a specially designed ABS sled. A GPS mast, centered over the two coils, completes the sled design. A 14-foot aluminum boat, powered by a gasoline outboard motor with a specialized propeller, tows the sled. A rope that is looped around the bow of the boat and attached at two points on the sled connects the two components. Poles mounted on the stern of the boat that extend below the waterline prevent the rope from being caught by the outboard motor propeller.

Figure 2-6. XTECH surveying in the open water area.

2.5.3 Data Processing Description

a. The geophysical data are reviewed on-site using Geosoft Oasis Montaj® software to ensure adequate survey coverage. Following this review, the data were electronically transferred to NAEVA's Virginia office for analysis/target selection.

b. Geosoft Oasis Montaj® UXO software package post-processed and contoured the raw data and identified potential UXO targets. The program identified peak amplitude responses of the frequency associated with, but not limited to, UXO items.

c. Geophysical data processing includes the following:

(1) Instrument drift correction (leveling).

(2) Lag correction.

(3) Digital filtering and enhancement (if necessary).

(4) Gridding of data.

(5) Selection of all anomalies.

(6) Selection of targets for intrusive characterization.

(7) Preparation of geophysical and target maps.

2.5.4 ATC Survey Comments

a. Several design shortcomings of this system affect both safety and performance. The bottom of the sled is a rectangular-shaped platform constructed with acrylonitrile butadiene styrene (ABS) pipe. Three pipes ride on the pond bottom parallel to the direction of boat travel; two are dragged perpendicular to the direction of travel (beneath the electromagnetic (EM) coils). At the front of the sled, the parallel pipes angle upward with a perpendicular support (fig. 2-7). This configuration permits objects to enter and then become trapped in the front of the sled.

Figure 2-7. Side view of sled.

b. The vertical component (GPS mast) of this sled is approximately 10 feet high and is also constructed of ABS pipe. An additional 4 feet of pipe were added when this system surveyed the deeper water area of the site (white pipe visible in fig. 2-6). Cement-filled pipes, placed inside the sled runners, served two purposes: to ensure the sled remained on the bottom and to lower the center of gravity (fig. 2-8).

Figure 2-8. Cement ballast.

c. This design did not work at this test site. The platform was unstable, particularly when turning. A second person, in a kayak, was needed to reorient the sled in an upright position. The amount of ballast placed in the ABS pipes along with the sled design dislodged an undetermined number of test items that had been emplaced either on the pond bottom or pressed into the sediment to be flush with the bottom.

d. At an actual MEC remediation site, this type of system would increase the chances of an explosive event. A 2-meter water depth and the length of the towrope will provide a limited level of personal protection, depending on the explosive item, but equipment replacement could be costly. Moving the GT items distorts the EM signatures, making item identification more difficult.

e. Overall, the towrope system did not work well. The system worked when pulling the sled in a straight line and in places where the boat made wide turns. However, when there is no tension on the towrope and the boat is maneuvering, the rope can become tangled in the propeller.

2.6 NAEVA GEOPHYSICAL, INC. (3DGEO)

2.6.1 Detection Instrumentation

The EM61 High Power (EM61 HP) system uses approximately 300 watts of transmit power instead of the approximately 100 watts in the standard system. The transmit waveform is bipolar instead of monopolar (current is driven one way and then the other and stacked). The transmit frequency of the high-powered transmitter is doubled when compared with a standard system. The net result of these improvements is to increase the transmitter moment from about 150 to 1200 amperes per square meter. Thus, the signal is increased, improving the signal-to-noise ratio of the recorded data. The effect almost doubles the recordable signal from any given target at a detectable depth. The system's depth of penetration is also increased.

2.6.2 Survey Platform

a. 3Dgeo's design for the underwater system incorporates three EM61 HP receiver coils and a single transmitter coil mounted on a carrying trailer made of rugged plastic sheets with structural separators and small stainless steel bolts. The design incorporates a simple skid between the wheels on the undercarriage of the cart to allow the trailer to skid over rough terrain or simply wheel over even ground. The design and construction of the platform allow it to work on land as well as submerged under as much as 15 feet of water (only evaluated in water depths of 8 feet or less).

b. NAEVA did not provide a description of the towing vessel. ATC's description is in section 2.6.5.

Figure 2-9. NAEVA/3Dgeo's sled and the towing vessel.

2.6.3 Data Processing Description

a. The geophysical data are reviewed on-site using Geosoft Oasis Montaj® software to ensure adequate survey coverage. Following this review, the data were electronically transferred to NAEVA's Virginia office for analysis/target selection.

b. Geosoft Oasis Montaj® UXO software package post-processed and contoured the raw data and identified potential UXO targets. The program identified peak amplitude responses of the frequency associated with, but not limited to, UXO items.

c. Geophysical data processing includes the following:

(1) Instrument drift correction (leveling).

(2) Lag correction.

(3) Digital filtering and enhancement (if necessary).

(4) Gridding of data.

(5) Selection of all anomalies.

(6) Selection of targets for intrusive characterization.

(7) Preparation of geophysical and target maps.

2.6.4 ATC Survey Comments

a. The towing vessel used both a gasoline outboard motor at the stern of the boat and two electric trolling motors, mounted to the port and starboard sides near the bow, for propulsion and maneuvering. The outboard motor provided the power needed to tow the sled along the bottom of the pond, while the thrust produced by the trolling motors helped to maneuver the boat into position for the next survey line. The electric motors also helped counteract some of the wind and wave actions that would otherwise force the boat off the required survey heading. Experimenting using both the forward and reverse thrust from just one trolling motor led to the elimination of the second unit.

b. The design of the bottom-riding sled allows it to maneuver easily along the contours that form the shoreline and in the open water at the center of the pond. The sled rests on four wheels (two swivel and two fixed) and connects to the boat by means of a rigid pole. The combination of motors on the towing vessel, the rigid pole, and swivel wheels allows the sled to make pivot turns. Aerodynamic design elements incorporated into the plastic sandwich body add to the stability and towability of the sled in water.

c. Overall, the design of this system makes it very maneuverable in a shallow water environment.

d. Prior damage to the GT targets in the open and deeper water areas limited this system to surveying the blind grid and littoral areas only.

2.7 AMEC EARTH AND ENVIRONMENTAL

2.7.1 Detection Instrumentation

The underwater towed sensor array consists of three Geometrics, Inc., model G882 mini-marine, cesium vapor magnetometers, a digital data recorder, and batteries. The G882 magnetometers record data with an absolute accuracy of < 3 nT at a rate of up to 10 Hz.

2.7.2 Survey Platform

a. The survey platform sled and towing vessel are the same as described in paragraphs 2.6.2 and 2.6.5 (fig. 2-10). The G882 magnetometers replaced the EM61 coils used in the preceding survey.

Figure 2-10. AMEC/3Dgeo towing platform with attached magnetometers.

2.7.3 Data Processing Description

a. During data collection, the magnetic and DGPS data are stored on the infield data logging system. At the completion of each day's survey, the data are downloaded to a laptop computer. MagMap2000® and MagPick®, magnetics processing software (Geometrics, Inc.), are used to preliminarily process the acquired data and review it for quality control. The DGPS positioning data are reviewed to make certain that data coverage gaps are not prevalent in the datasets.

b. MagMap2000® and MagPick® software analyzed the survey data a second time to generate the dig list. MagPick® generates estimates of the X, Y, Z position and mass of magnetic anomalies. It implements a geophysical inversion technique using maximum likelihood dipole pattern matching methods to analyze anomalies. In this approach, the basic nature of the anomaly source is considered to be known (point object or sphere), and then to check the quality of the model, a synthetic magnetic field is calculated from the model and compared with the observed one. UXO targets produce a magnetic field distortion that can be accurately approximated with the well-defined fields of magnetic dipoles or uniformly magnetized spheres. The magnetic dipole itself is characterized by six unknown values: X, Y, and Z as coordinates of its center and Jx, Jy, Jz, which represent values of the magnetic moment. The mathematical inversion task performed by MagPick® is summarized as given magnetic field T observed in the vicinity of the object; the unknown values X, Y, Z, Jx, Jy, and Jz are varied such that the computed field C assumes maximum similarity with the observed field T observed.

c. Identification of the magnetic anomalies uses both manual and automated methods on the gridded datasets. An anomaly-modeling algorithm in the MagPick® software is used to determine the total magnetic moment (cgs), burial depth (m), and mass (kg) of each interpreted anomaly. Class identification and discrimination of the magnetic anomalies were achieved by comparing the magnetic moment and mass of the known UXO in the calibration lane to the modeled moment and mass of the magnetic anomalies.

d. The 60-mm UXO in the calibration lane did not produce a magnetic anomaly. Therefore, an interpretation of the 60-mm UXO is not contained in the datasets.

2.7.4 ATC Survey Comments

a. The towing vessel used both a gasoline outboard motor at the stern of the boat and two electric trolling motors, mounted to the port and starboard sides near the bow, for propulsion and maneuvering. The outboard motor provided the power needed to tow the sled along the bottom of the pond, while the thrust produced by the trolling motors helped to maneuver the boat into position for the next survey line. The electric motors also helped counteract some of the wind and wave actions that would otherwise force the boat off the required survey heading. Experimenting using both the forward and reverse thrust from just one trolling motor led to the elimination of the second unit.

b. The design of the bottom-riding sled allows it to maneuver easily along the contours that form the shoreline and in the open water at the center of the pond. The sled rests on four wheels (two swivel and two fixed) and connects to the boat by means of a rigid pole. The combination of motors on the towing vessel, the rigid pole, and swivel wheels allows the sled to make pivot turns. Aerodynamic design elements incorporated into the plastic sandwich body add to the stability and towability of the sled in water.

c. Overall, the design of this system makes it very maneuverable in a shallow water environment.

d. Prior damage to the GT targets in the open and deeper water areas limited this system to surveying the blind grid and littoral areas only.

SECTION 3. SYSTEM MANEUVERABILITY EVALUATION

The first challenge in surveying a shallow water site is to traverse the entire area with the sensing system. The comments in paragraphs 2.X.4 of each subtest discuss the maneuverability, strengths, and weaknesses for each system based on field observations made during site surveys. This section quantifies the percentage of each test area surveyed based on ATC's post-survey analysis of the demonstrator's data. The relationship between the percentage of area surveyed and system scoring is also discussed.

3.1 TEST ENVIRONMENT AND PRESURVEY INFORMATION

a. The design of the shallow water site presents a limited number of navigational and maneuvering challenges to the surveyors. These challenges fall into four general categories: topography (shorelines, bottom contours) submerged obstacles (man-made and naturally occurring); vegetation; and atmospheric (winds and associated wave action). These challenges are very mild compared with what can be encountered in marine environments.

b. Each demonstrator claimed his system could survey all of the test areas within the shallow water site.

c. Prior to arriving at the test site, each demonstrator received the following:

(1) A spreadsheet entitled "Area Boundaries," which contained sets of coordinates (minimum 7, maximum 34) that formed a polygon around each test area. The polygons intentionally overlapped the adjoining test areas to ensure complete coverage of the individual test area (the automated scoring routine assigns any identified object to the correct test area).

(2) A brief description of each area.

(3) The GPS coordinates for 96 items in the pond (81 shot puts and 15 calibration targets).

(4) The vertical and horizontal data for each of three first-order survey monuments near the test area as well as a permanent control point adjacent to the pond.

d. Each system surveyed the pond in any manner that the operators felt was appropriate. The surveys ended when the demonstrators were satisfied with the data they had collected.

3.2 SURVEYED AREA CALCULATION

The percent coverage for a given test area is determined by first plotting the raw GPS coordinates combined with the sensor swath (line spacing and associated overlap), calculating the area surveyed, and then comparing the surveyed area with the total test area.

$$\frac{\text{Section Surveyed}}{\text{Test Area Size}} \times 100 = \text{\% Surveyed}$$

3.3 AREA SURVEYED AND SCORING RELATIONSHIP

a. Both the test and scoring methodologies required the demonstrator to survey 100 percent of each of the four test areas (blind grid, open water, littoral, and deeper water). Scoring a partially surveyed area alters the ordnance and clutter sample sizes, changes the test area boundaries, and decreases the statistical confidence in the performance statements made for that area. Allowing partial scoring decreases the validity of performance comparisons made between multiple test areas for a single demonstrator and comparisons made between multiple demonstrators for a single test area.

b. The demonstrator's system is always scored against the complete GT for a given test area regardless of the percentage of the area covered.

c. With the realization that some systems may not be able to survey 100 percent of a given test area, a ranking system was established. The ranking system description and the area-surveyed results for all areas and demonstrators are in Table 3-1. The ranking system was established in support of a chi-square 2 x 2 contingency test used for comparing system performance across test areas in the individual system reports. That analysis tool is not used in this report.

3.4 RESULTS

TABLE 3-1. SURVEY RANKING SYSTEM AND AREA RESULTS

Test Area	Geophex	Tetra Tech	CTC Foerster	NAEVA XTECH	NAEVA 3Dgeo	AMEC 3Dgeo	IT Jewell G882	IT Jewell MM
Blind grid	100	100	100	98	100	100	78	81
Open water	99	99	84	84			69	58
Littoral	82	88	74	10	[a]84	[a]80	64	70
Deeper water	98	98	94	65			72	58

% Area Covered	Ranking	
95 to 100	Met	Data used for the direct comparison between systems and areas.
90 to 94	Generally met	Data used for a comparison between systems and areas. A small negative bias is contained in the reported numbers (bias not quantified in this report).
50 to 89	Partially met	Data reported but not compared between systems or areas. A large negative bias is contained in the reported numbers (bias not quantified in this report).
0 to 49	Not met	Data not scored or reported.
	Unavailable	Test area not available for surveying.

[a]The time interval between the first five surveys and this one allowed phragmites to overgrow part of the littoral region that had been seeded with GT targets. This presented more of a challenge for these systems.

3.5 ANALYSIS

a. All of the tested systems achieved their highest percent of area coverage in the blind grid area. By design, this area was free of all navigational challenges, easily accessible from all sides, uniform in water depth, and required a system response at each of 644 given locations (i.e., a very controlled, homogenous environment). The percentage of surveyed area decreases as each system encounters the challenges in the other test areas.

b. Based on field observations made during the surveys and the post-test review of the submitted data, the reason some systems did not meet the area coverage requirement is a combination of the following:

(1) Limitations of the sensor package or survey platform design - it could not physically get into an area.

(2) GPS limitations - resolution on line spacing or positioning errors.

(3) Poor quality control/quality assurance measures.

SECTION 4. SCORING PROCEDURES

a. ATC's analysis of the data from the shallow water site in this report differs from the standardized scoring procedures developed by USAEC for two reasons. The first is that an analysis, based on the USAEC procedures, for each system described in this document is already available on the World Wide Web at http://aec.army.mil (the scoring procedures and data summaries from those individual reports are also in app A and B respectively). The second is to provide additional information on system performance that will help both the detection system designers/operators and remediation site managers gain a better understanding of the current system capabilities.

b. There are two key differences between the USAEC analytical procedures and this one. The first is that clutter is tracked and reported in both the response and discrimination stages. The second is that all analyses use the discriminated target declarations provided by the system demonstrators in their final data submission package (i.e., their dig lists). The percentage of items detected by each system in this report agrees to within a few points of the values obtained using the USAEC procedures. The differences are caused by the reclassification of targets that sometimes occurs as the receiver operating characteristics (ROC) curve rules are applied to the demonstrator's dig list.

c. This report presents the results of the response and discrimination analysis in both a summarized and a detailed format.

d. All of the calculated values in this report assume that the number of detections is a binomially distributed random variable. Reported results are at the 90-percent reliability, 95-percent confidence levels unless otherwise noted.

4.1 SUMMARY RESPONSE SCORING

Central to each shallow water system is a metal detector and its associated data collection/recording devices. The first part of this analysis looks at how well each system can detect metal. There are two terms associated with this analysis: probability of metal detection P_m^{res} and the probability of false metal detection P_{fm}^{res}. Both of these terms are defined in Figure 4-1.

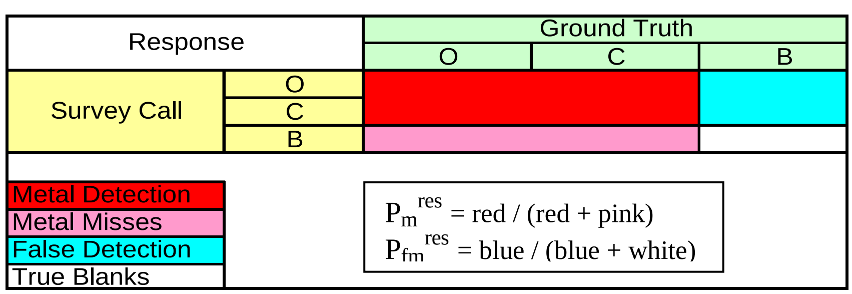


Figure 4-1. Response definitions.

4.2 SUMMARY DISCRIMINATION SCORING

a. The metal objects detected by these systems fall into one of two general GT categories, ordnance or clutter. This part of the analysis looks at how well each system can discriminate between these general categories. There are two terms associated with this analysis: probability of ordnance discrimination P_o^{dis} and probability of clutter discrimination P_c^{dis}. Both of these terms are defined in Figure 4-2.

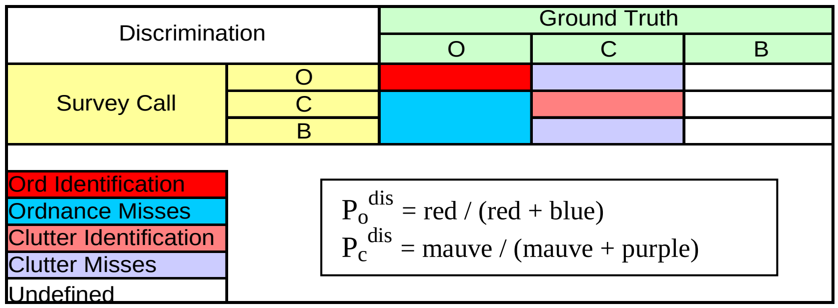


Figure 4-2. Ordnance and clutter identification definitions.

b. The four values from the summary calculations are presented as comparative bar graphs showing the results of all demonstrators for a particular test area. This provides a general overview of each system's performance.

4.3 DETAILED RESPONSE AND DISCRIMINATION ANALYSIS

a. The difference between the preceding analyses and this one is the baseline to which the demonstrator's results are compared. In the preceding analyses, the response and discrimination probabilities are based on the total number of emplaced items, existing blank cells in the blind grid, or a percentage of open space in the other test areas. In this analysis, each detected item is associated with a specific GT category and reported as a percentage of the number of items contained within that particular category. This provides the assessment of detection ability by item type. Then the numbers of items, within each category, correctly identified as either ordnance or clutter, are counted. This provides the discrimination capability by item type. The results are reported as a percentage of items emplaced to protect the GT.

b. This analysis provides a graphical presentation of the strengths and weaknesses of each system tested. An explanation of the color codes used in the graphs and their significance is given in section 5.2.4. This places the explanation closer to the first graph to which it applies.

SECTION 5. SYSTEM PERFORMANCE BY TEST AREA

The design of each test area within the shallow water site provides unique data on the system's capability. This section provides additional information on the design of each test area, explains the purpose of the data collected, and compares all of the demonstrated systems within the given area.

5.1 CALIBRATION TEST GRID

5.1.1 Description

This area contains 15 projectiles, 3 each 40, 60, 81, 105, and 155 mm. One of each projectile type is buried at the projectile diameter to a depth ratio shown in Figure 5-1. In addition, two clutter-cloud target scenarios were constructed adjacent to this area. The varying size of the cells and the burial depths of the projectiles within the cells minimized signal overlap.

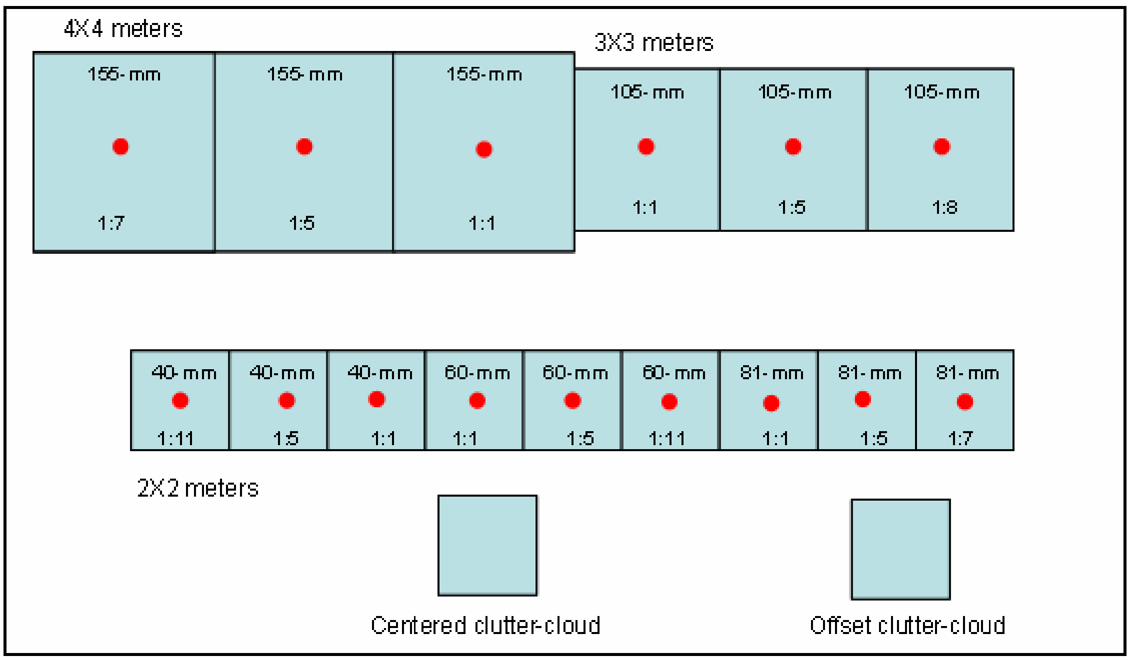


Figure 5-1. Schematic of the calibration grid.

5.1.2 Purpose

The calibration grid provides the demonstrators with an opportunity to measure their instrument's responses over ordnance items at known locations and emplacement depths. This provides the user with a sensor library of detection responses for the emplaced targets and an understanding of their resistivity prior to entering the blind test fields.

5.1.3 Analysis

The data collected in this area is for the demonstrator's internal use only. System performance is not reported using this area.

5.2 BLIND TEST GRID

5.2.1 Description

a. This area provides a statistically valid (90-percent reliability/95-percent confidence levels), controlled environment in which the demonstrator must provide a response (ordnance, clutter, or blank) at each of the 644 grid locations. Each grid cell is 2 by 2 m^2. At the center of each cell is either an ordnance item, clutter, or nothing. Surrounding the blind grid on three sides are 3.6-kg (8-lb) shot puts, buried 0.3 meter deep in the sediment (fig. 5-2). The shot puts define the grid boundaries and can be used as a navigational/GPS check. The GPS coordinates for the center of each cell and the shot put locations were given to each surveyor prior to testing.

b. All of the GT items in this area are located under 8 feet of water and 1 foot (measured to the top of the emplaced item) of soil.

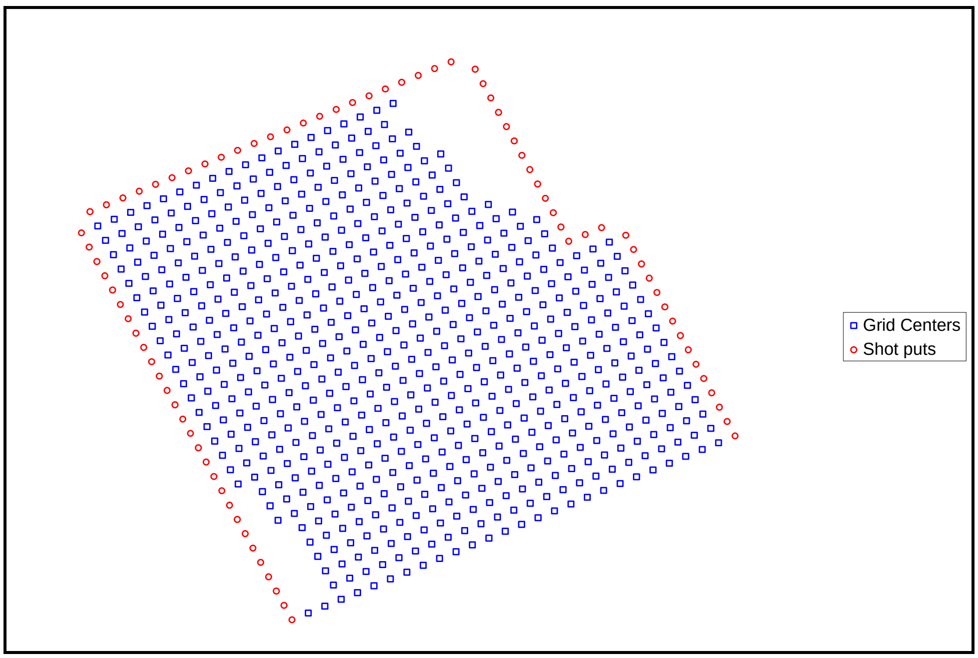


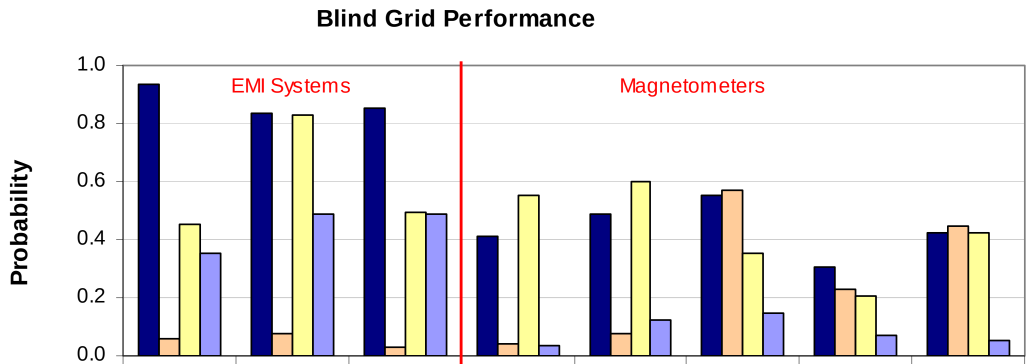
Figure 5-2. Blind grid schematic.

5.2.2 Purpose

The best indication of system performance is in the blind grid. Comparison of the demonstrator's response and discrimination lists to the GT in this area determines the range of ordnance the system can reliably detect and establishes the baseline against which system performance in the other test areas is measured.

5.2.3 Response and Discrimination Findings

a. The performance of the various demonstrators against the GT based on the terms and definitions in section 4 is shown in Figure 5-3.



	NAEVA / XTECH	Geophex	NAEVA / 3Dgeo	CTC	AMEC /3Dgeo	ITJ MM	ITJ 882	TetraTech
Metal Detection	0.93	0.84	0.86	0.41	0.49	0.55	0.30	0.42
False Metal Detection	0.06	0.08	0.03	0.04	0.07	0.57	0.23	0.45
Ord Correctly ID	0.46	0.83	0.50	0.55	0.60	0.35	0.21	0.42
Clutter Correctly ID	0.35	0.49	0.49	0.03	0.13	0.15	0.07	0.05

Figure 5-3. Blind grid surveyor performance summary.

b. There are several points to keep in mind when viewing all of the graphs in this report:

(1) The clutter in each site is composed of both ferrous and nonferrous material.

(2) Some surveyors used magnetometers while others used EM sensors. Dual-mode (magnetometer and EM) systems were not tested.

(3) All systems are scored using the same GT (impact areas contain both ferrous and nonferrous of clutter).

(4) All systems are scored against the complete GT in a test area regardless of the percentage of the area surveyed.

(5) The choice of sensor influences the apparent GT in both the detecting and discrimination processes. In particular, the total number of metal objects detected, or conversely, the amount of blank space.

5.2.4 Detailed Findings

a. The following is the explanation of the color codes and their significance in the detailed analysis process described in paragraph 4.3.

b. The sum of the blue and yellow bars represents the true percentage of items detected for each GT category. The red bars show the percentage of items missed by category. Together, this provides an indication of the strengths and weaknesses of the detection system.

c. The blue bars represent the percentage of each category correctly discriminated. In this case, the discrimination is only between ordnance and clutter, not individual caliber or clutter sizes. The yellow bars indicate the misidentified percentage of each class of item, that is, either an ordnance item called clutter or vice versa. Together, this provides an indication of the strengths and weaknesses of the discrimination system.

d. The stacked bar labeled "blank" is a unique category. The misidentified percentage in this column represents the quantity of blank (empty) cells identified as containing either ordnance or clutter.

e. These codes describe the GT categories in Figures 5-4 through 5-11: numeric values = the ordnance caliber in millimeters, F = ferrous clutter, N = nonferrous clutter, L = large, M = medium, and S = small.

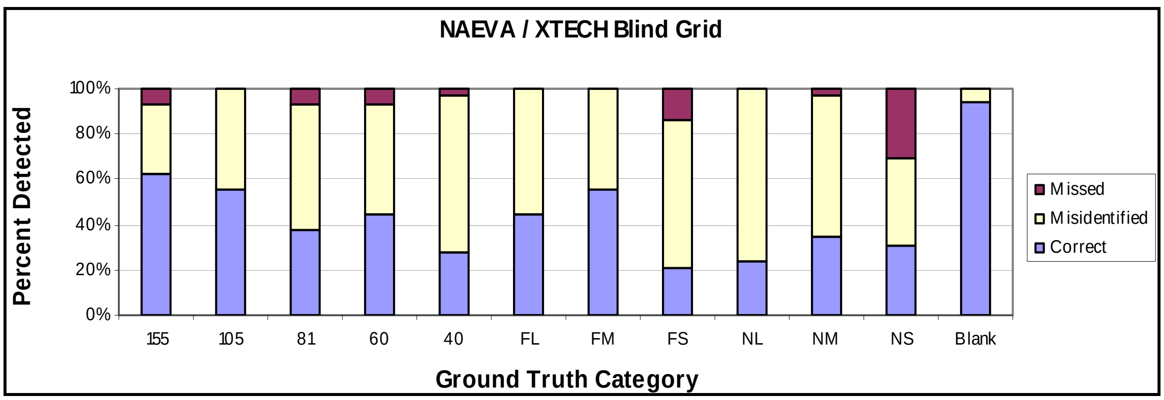


Figure 5-4. NAEVA/XTECH detection and discrimination results.

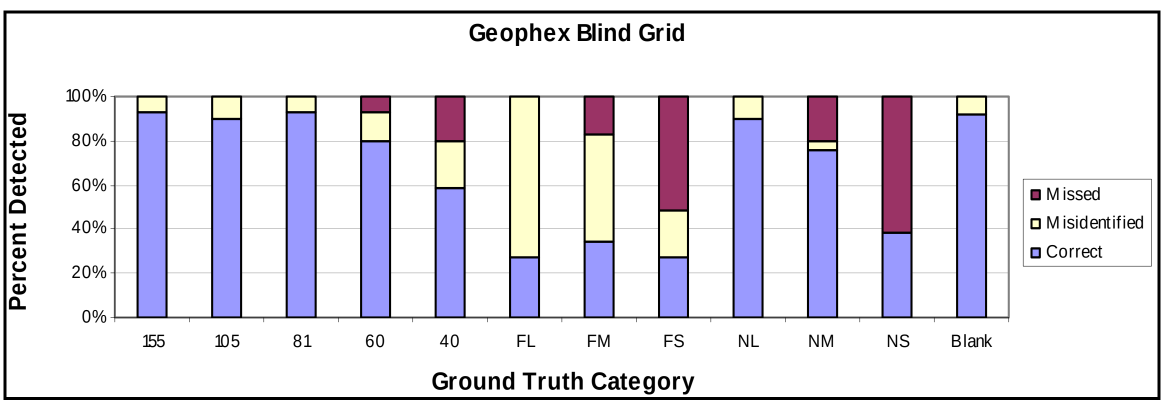


Figure 5-5. Geophex detection and discrimination results.

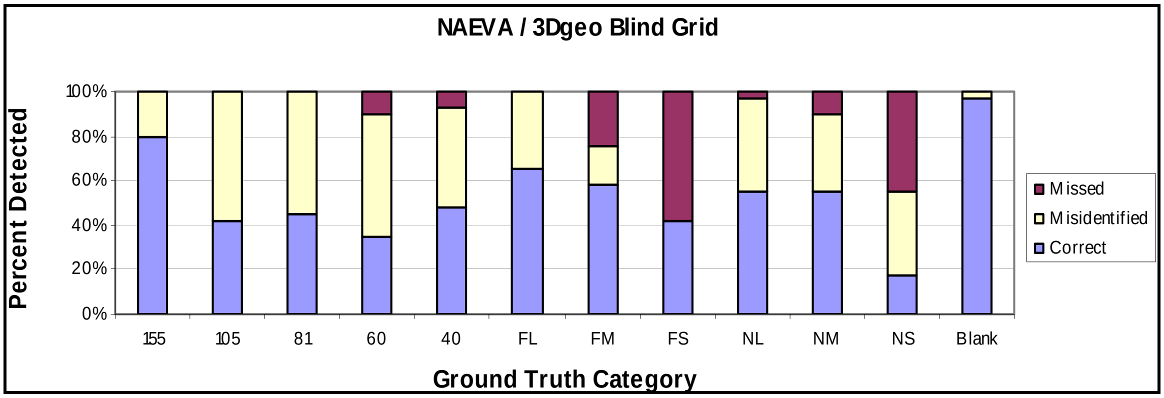


Figure 5-6. NAEVA/3Dgeo detection and discrimination results.

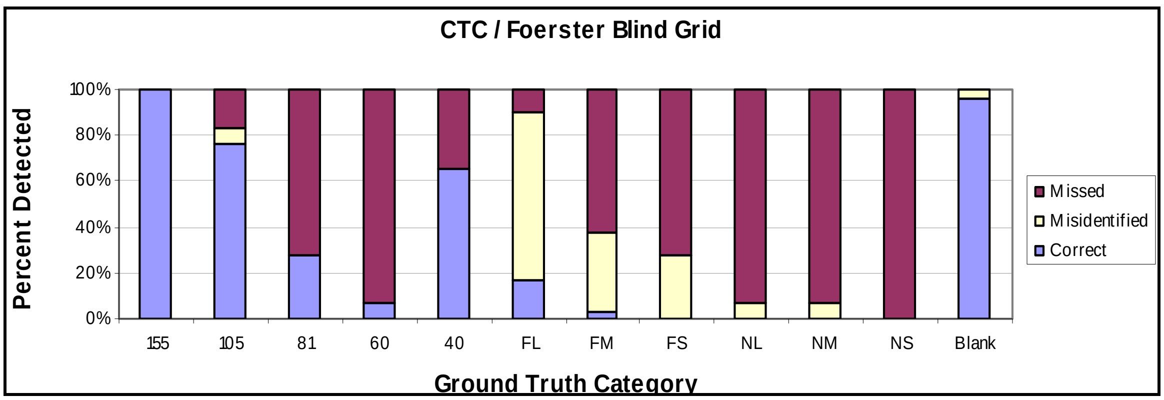


Figure 5-7. CTC/Foerster detection and discrimination results.

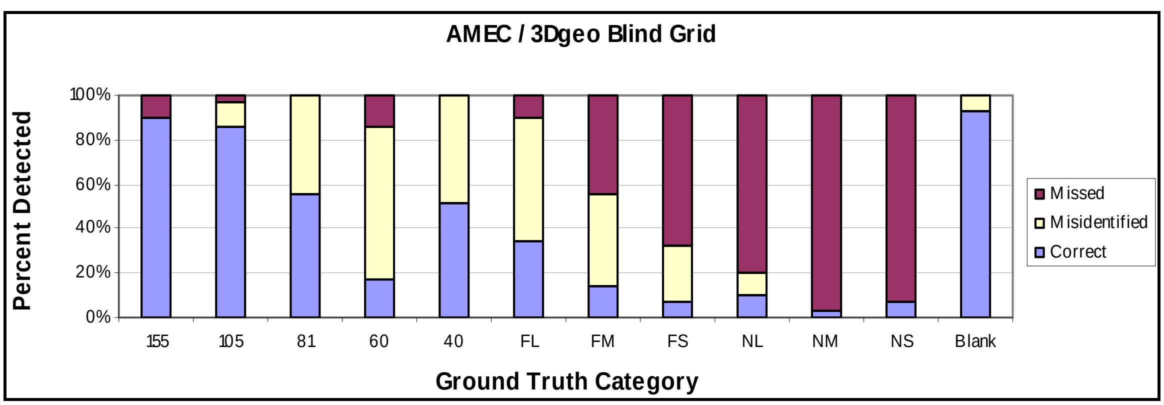


Figure 5-8. AMEC/3Dgeo detection and discrimination results.

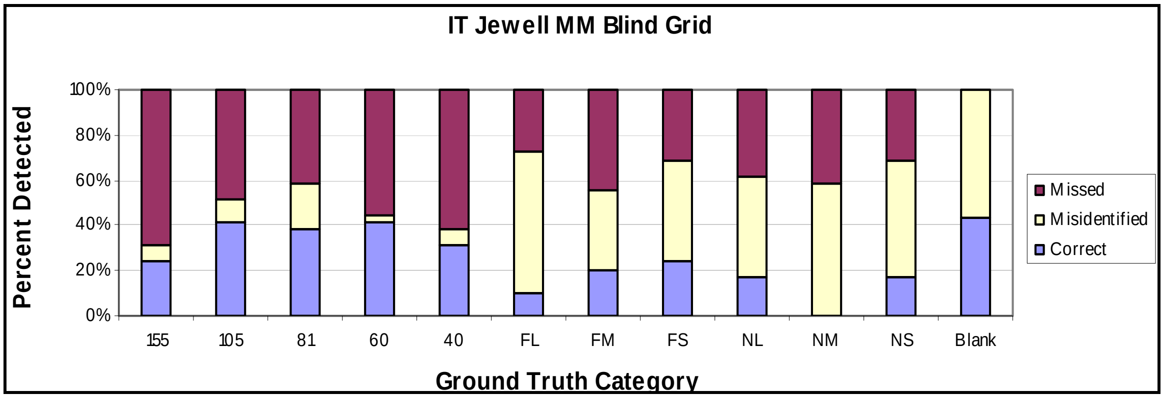


Figure 5-9. IT Jewell MM detection and discrimination results.

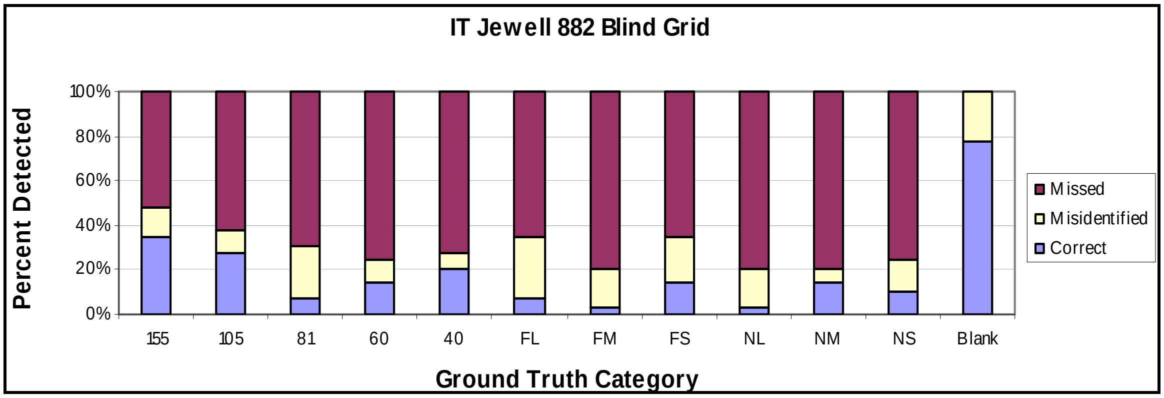


Figure 5-10. IT Jewell 882 detection and discrimination results.

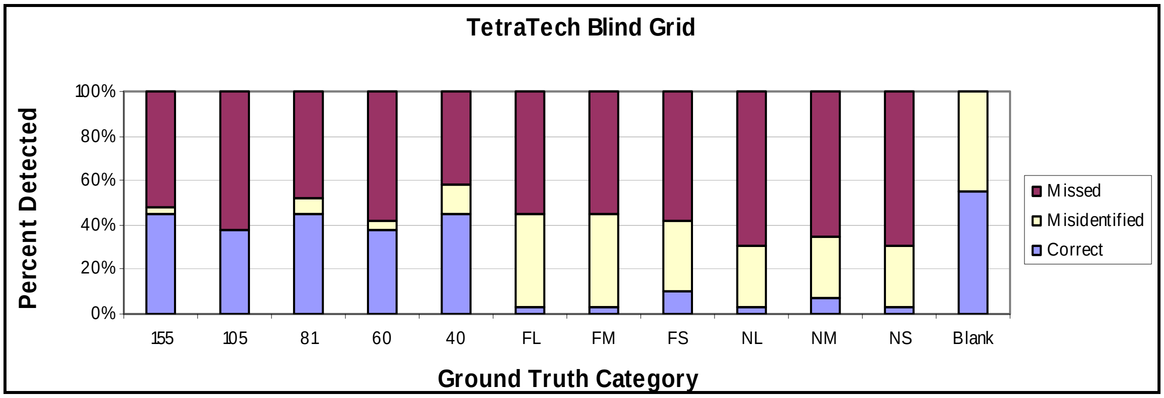


Figure 5-11. Tetra Tech detection and discrimination results.

5.2.5 Blind Grid Technical Assessment

a. The most obvious conclusion to be drawn from the chart in Figure 5-3 is that EMI sensors outperformed the magnetometer sensors. The EMI sensors detected a greater number of metal objects. To put that in perspective, approximately 27 percent of the GT clutter items is comprised of nonferrous materials that magnetometers typically do not find. Looking at the probability of identifying ordnance (ferrous) items allows a more balanced comparison between the sensor types. The EMI systems still identified more ordnance items than the magnetometers; however, the performance difference is smaller. The detailed graphs show the differences between all of the systems that surveyed the blind grid site.

b. Sensor deployment also contributed significantly to both the detection and the discrimination results. The water depth in the blind grid area is 8 feet; the target depth is 1 foot. The sensor-to-target distance for each detection system is shown in Table 5-1. A generally accepted, maximum distance limit for magnetometer detection/discrimination is 6.6 feet (2 m).

TABLE 5-1. SENSOR-TO-TARGET DISTANCE

Demonstrator	**Sensor Type**	**Deployment**	**Target Standoff [a]**	
			ft	**m**
Tetra Tech	Magnetometer	Surface	10	3.0
IT Jewell MM	Magnetometer	Subsurface	7.5	2.3
IT Jewell 882	Magnetometer	Subsurface	7.5	2.3
CTC/Foerster	Magnetometer	Near bottom	3.0	0.9
AMEC/3Dgeo	Magnetometer	Bottom riding	2.5	0.8
NAEVA/3Dgeo	EMI	Bottom riding	2.5	0.8
Geophex	EMI	Bottom riding	2.0	0.6
NAEVA/XTECH	EMI	Bottom riding	1.3	0.4

[a]Perpendicular distance between sensor and GT

c. The bottom-riding systems produced the highest detection rates followed by the adjustable depth magnetometer system. The surface riding and shallow deployment systems performed poorly based on the high probability of false metal detections. This reduced performance is also reflected in the detailed discrimination charts.

d. Distance between the sensor and an emplaced object is only one component of the differences in detection rates. The site survey plans submitted by the demonstrators and field observations made during the surveys showed variations in the way sensors were orientated, how the GPS systems were employed, sampling rates, survey speeds, line distances, and how stated quality control and assurance procedures were applied, or not applied. An assessment of the individual demonstrator’s procedures is outside the scope of this project.

e. The discrimination process is dependent on a response of sufficient resolution to make a distinction between objects. Each demonstrator has a unique discrimination process; however, all of the processes rely on a combination of computer-processed data and images and then on the human interpretation of that information. An assessment of the discrimination process is also outside the scope of this project.

5.3 OPEN WATER AREA

5.3.1 Description

a. The open water area contains a variety of navigational, detection, and discrimination challenges. It is 2.1 acres (8396 m^2) in size. Water depths range from 6 to 10 feet. This body of water is not subject to tides or currents.

b. The GT in this area contains, as a subset, all of the items found in the blind grid. An additional category of GT, called mixed metals (MX), along with 8-inch projectiles and a few clutter items not contained in the blind grid, has been added to this area. The GT burial depths have been randomized.

5.3.2 Purpose

This area provides several challenges not seen in the blind grid and is more representative of shallow water environments. As such, it provides a better indication of the detection and discrimination abilities of each system at an actual remediation site.

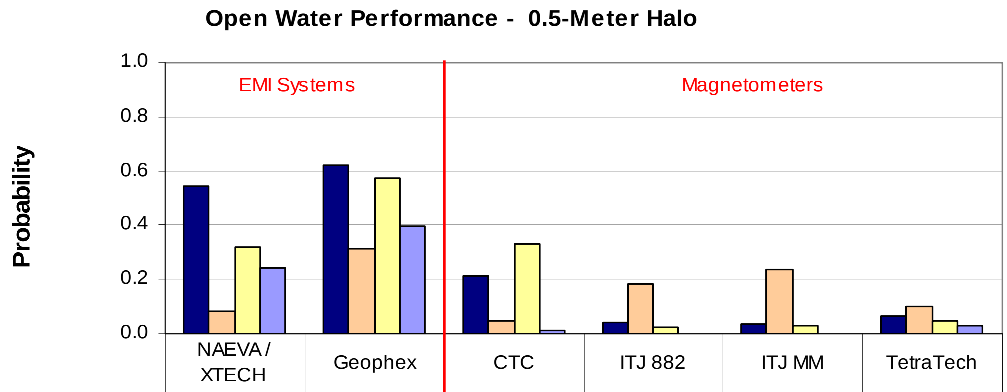
5.3.3 Response and Discrimination Findings

a. The calculated values in this section assume that the number of detections is a binomially distributed random variable. All results are at the lower 90-percent reliability/95-percent confidence levels unless otherwise noted.

b. Prior damage to the open water site prevented two of the demonstrators (NAEVA/3Dgeo and AMEC/3Dgeo) from surveying this area. Two of the six systems shown (Geophex and Tetra Tech) met the area coverage requirements. The other four surveyed 84 percent or less of the test area (table 3-1).

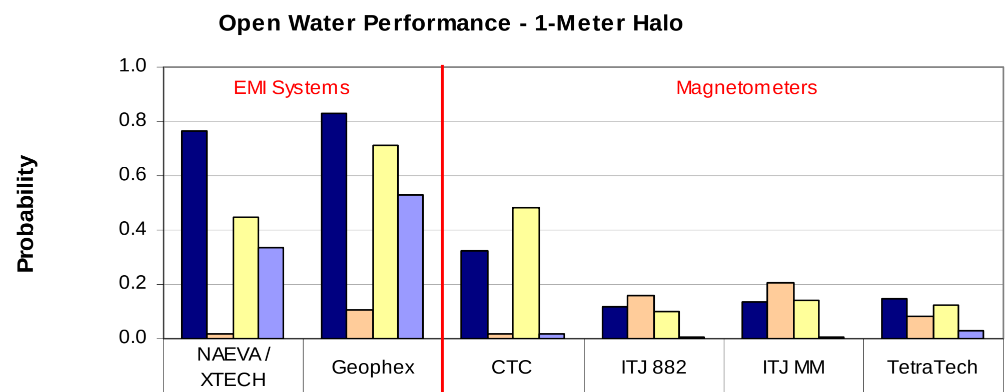
c. System performance was first evaluated using a 0.5-meter detection halo. Comparing the open water detection and discrimination results to those obtained in the blind grid at this halo size shows a marked decrease in the performance for all surveyors. The dig-sheet data were reanalyzed using a 1-meter halo.

d. The performance of the various demonstrators using both halo sizes based on the terms and definitions in section 4 is shown in Figures 5-12 and 5-13.



	NAEVA / XTECH	Geophex	CTC	ITJ 882	ITJ MM	TetraTech
Metal Detected	0.54	0.62	0.21	0.04	0.04	0.07
False Metal Detection	0.08	0.31	0.05	0.18	0.24	0.10
Ord Correctly Identified	0.32	0.57	0.33	0.03	0.03	0.04
Clutter Correctly Identified	0.24	0.39	0.01	0.00	0.00	0.03

Figure 5-12. Open water system performance, 0.5-meter halo.



	NAEVA / XTECH	Geophex	CTC	ITJ 882	ITJ MM	TetraTech
Metal Detected	0.77	0.83	0.33	0.12	0.14	0.15
False Metal Detection	0.02	0.10	0.02	0.16	0.21	0.08
Ord Correctly Identified	0.45	0.71	0.48	0.10	0.14	0.12
Clutter Correctly Identified	0.33	0.53	0.02	0.00	0.00	0.03

Figure 5-13. Open water system performance, 1.0-meter halo.

5.3.4 Detailed Findings

a. Two additional codes have been added to the GT category: 8-inch = 8-inch artillery projectiles and MX = mixed metals.

b. The detailed detection and discrimination results shown in Figures 5-14 through 5-19 are based on the procedures in section 4. The scatter graphs show the difference between the emplaced GT center and the surveyor-reported position for all true metal detections (ordnance and clutter). The positioning deltas are calculated using the square root of the sum of the squares of the northing and easting errors.

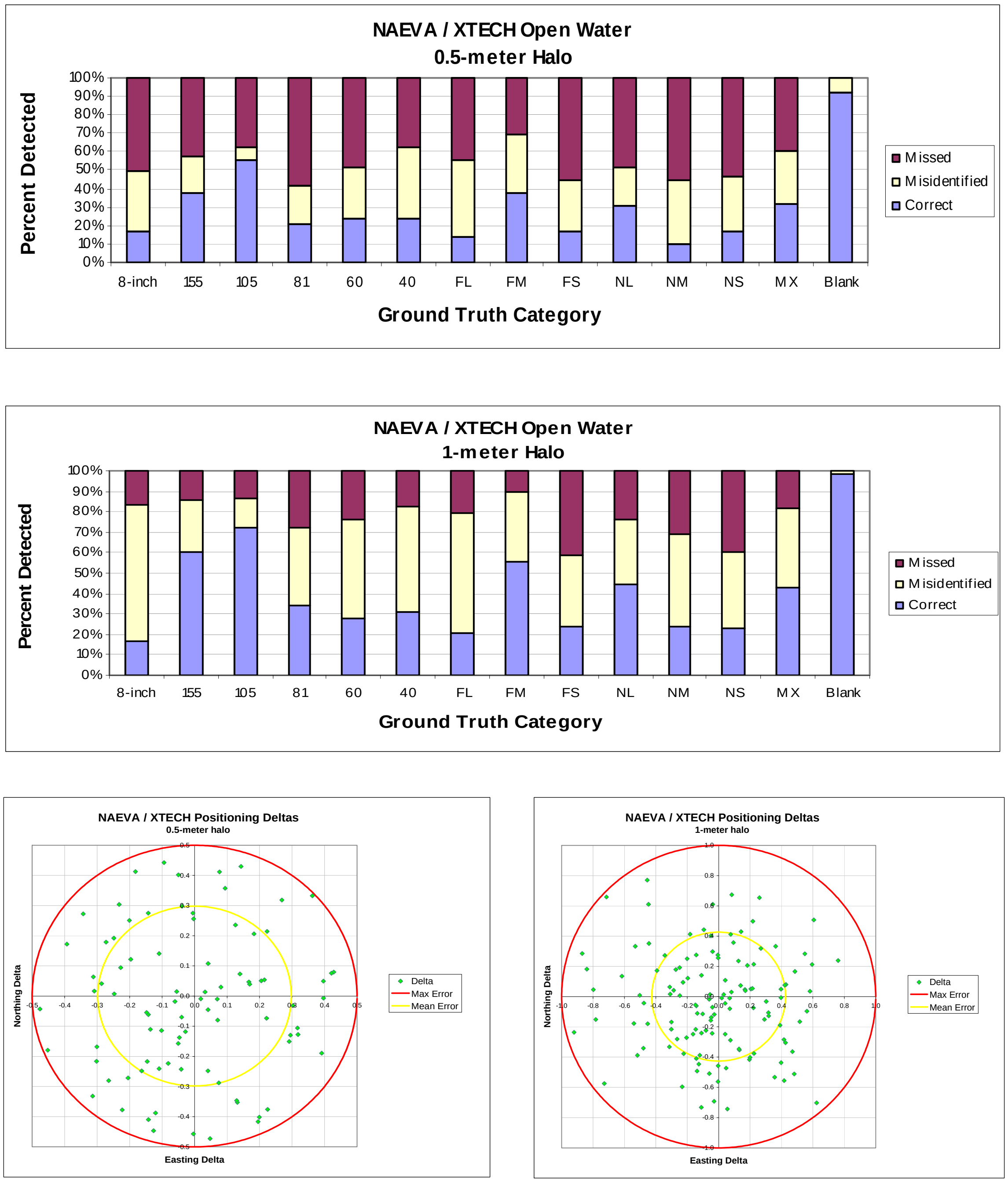


Figure 5-14. NAEVA/XTECH open water performance.

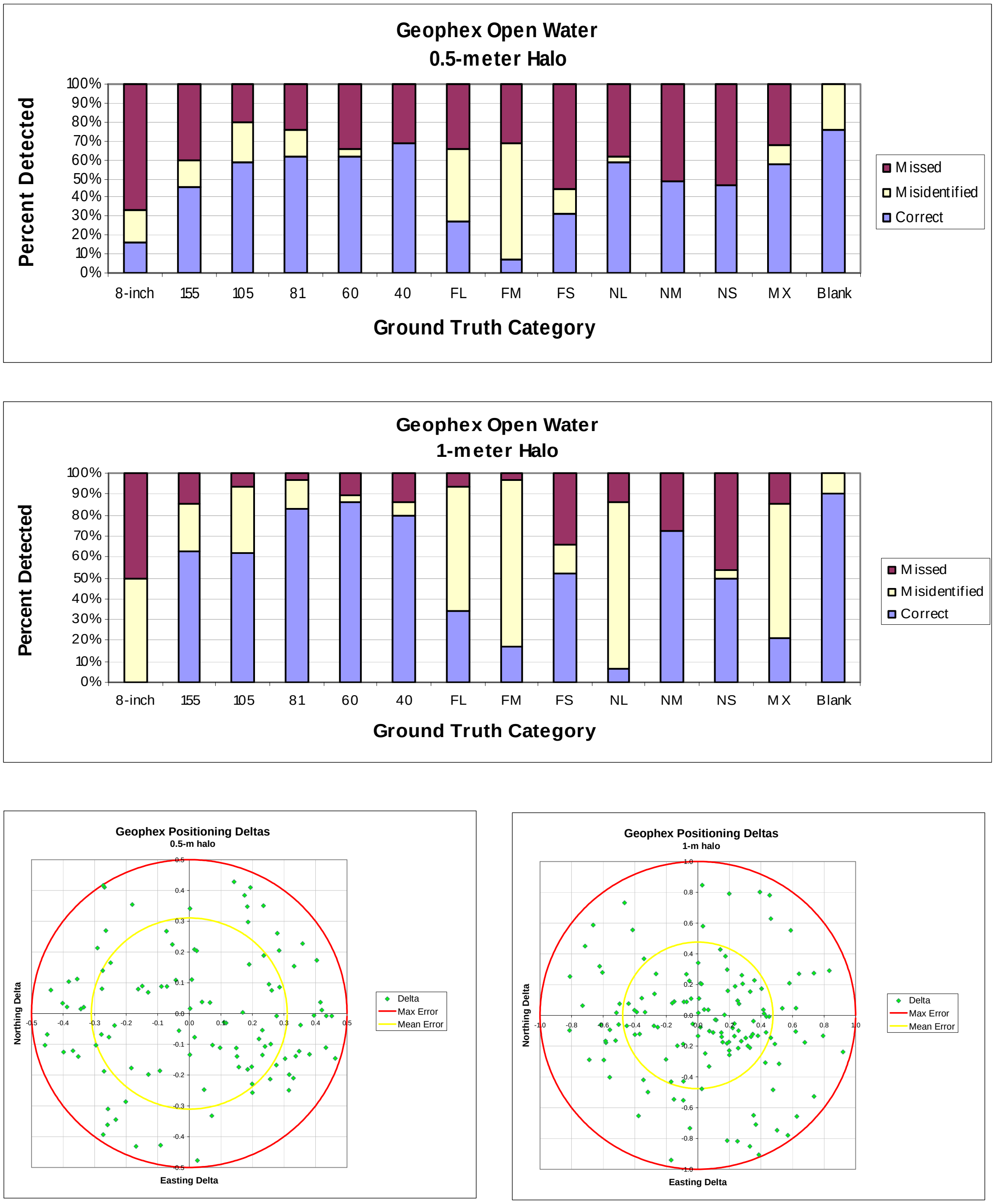


Figure 5-15. Geophex open water performance.

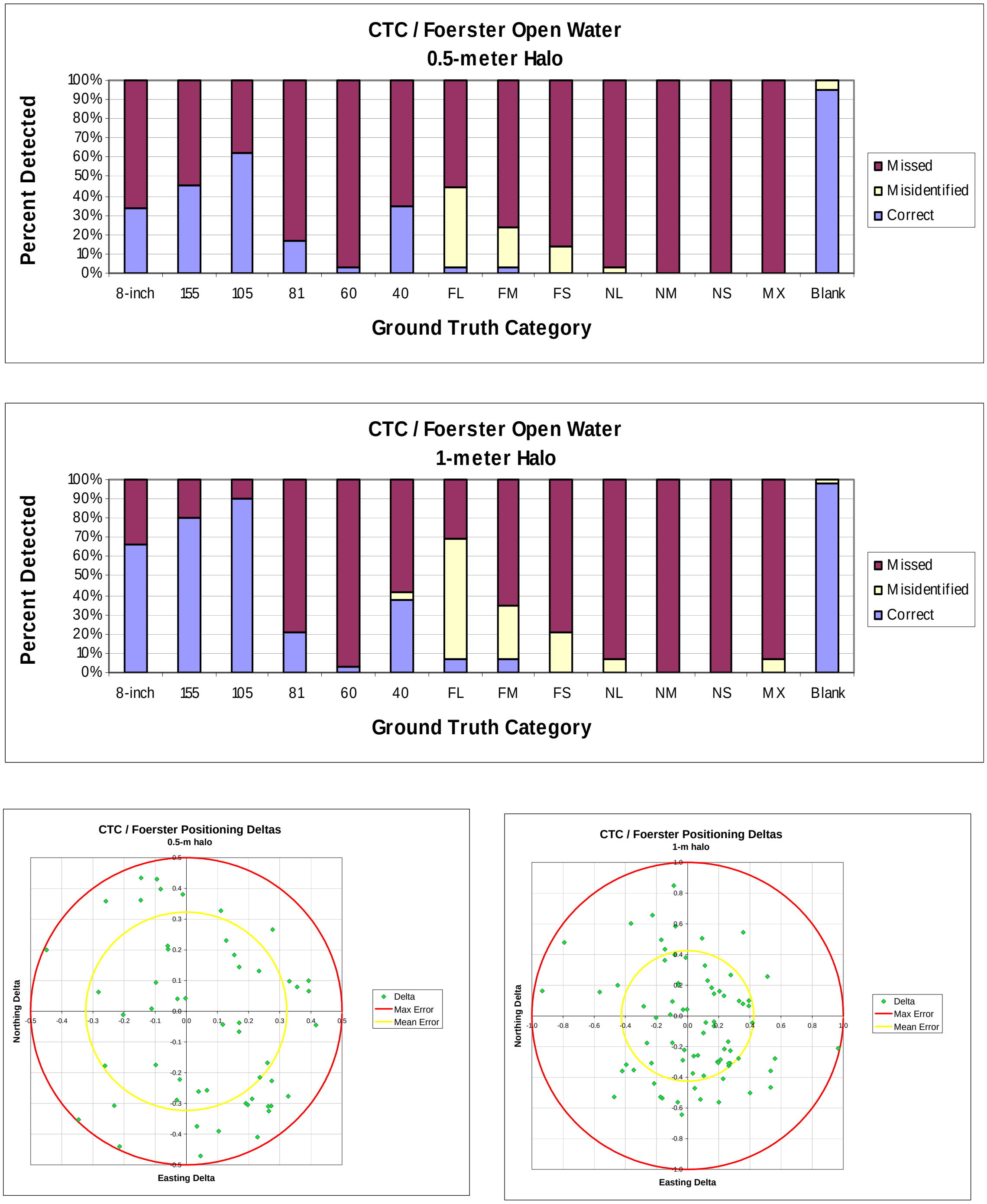
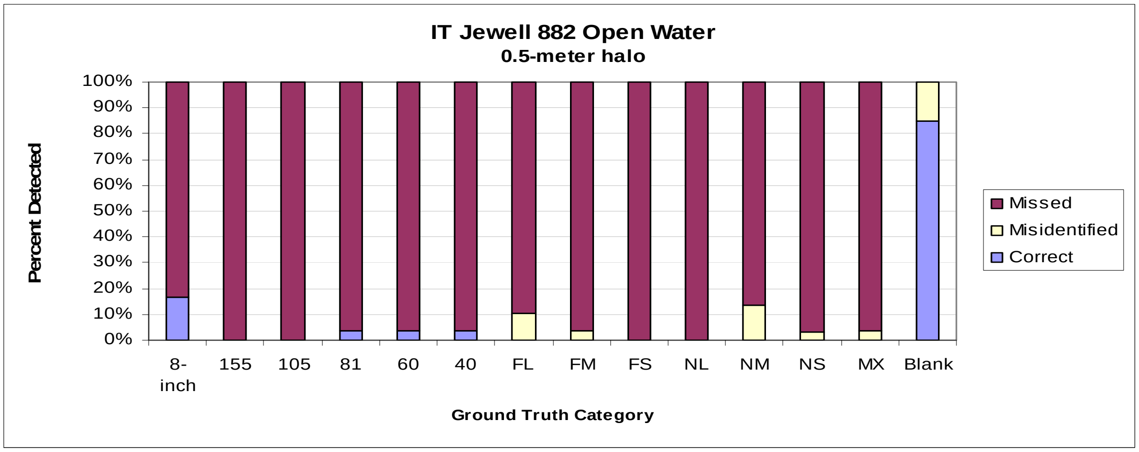
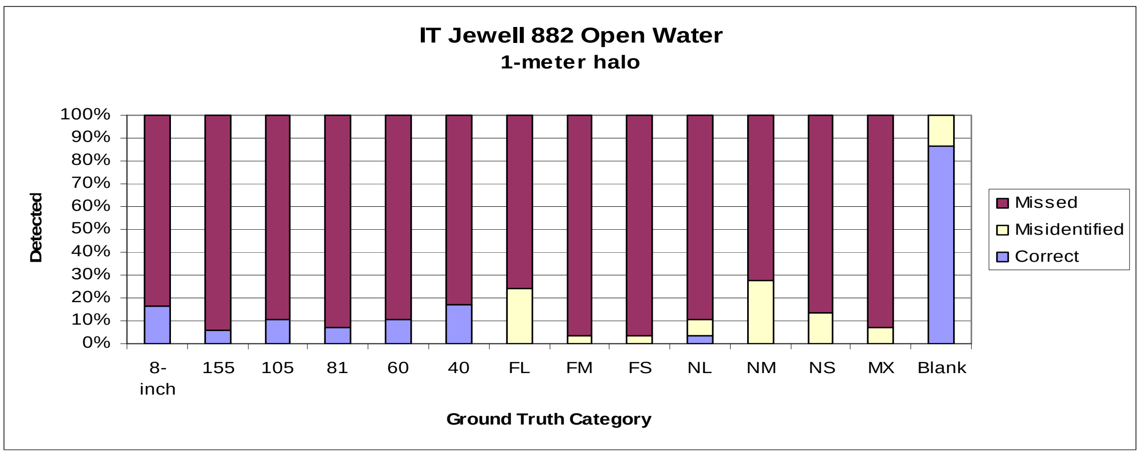
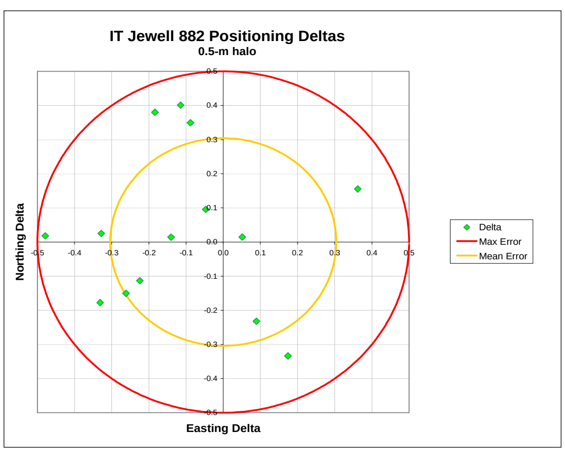


Figure 5-16. CTC/Foerster open water performance.







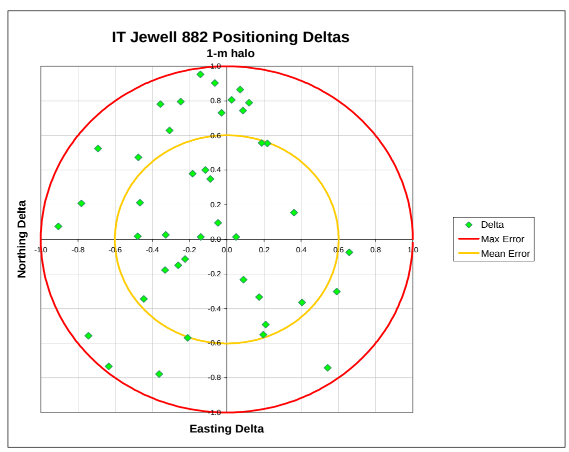


Figure 5-17. IT Jewell 882 open water performance.

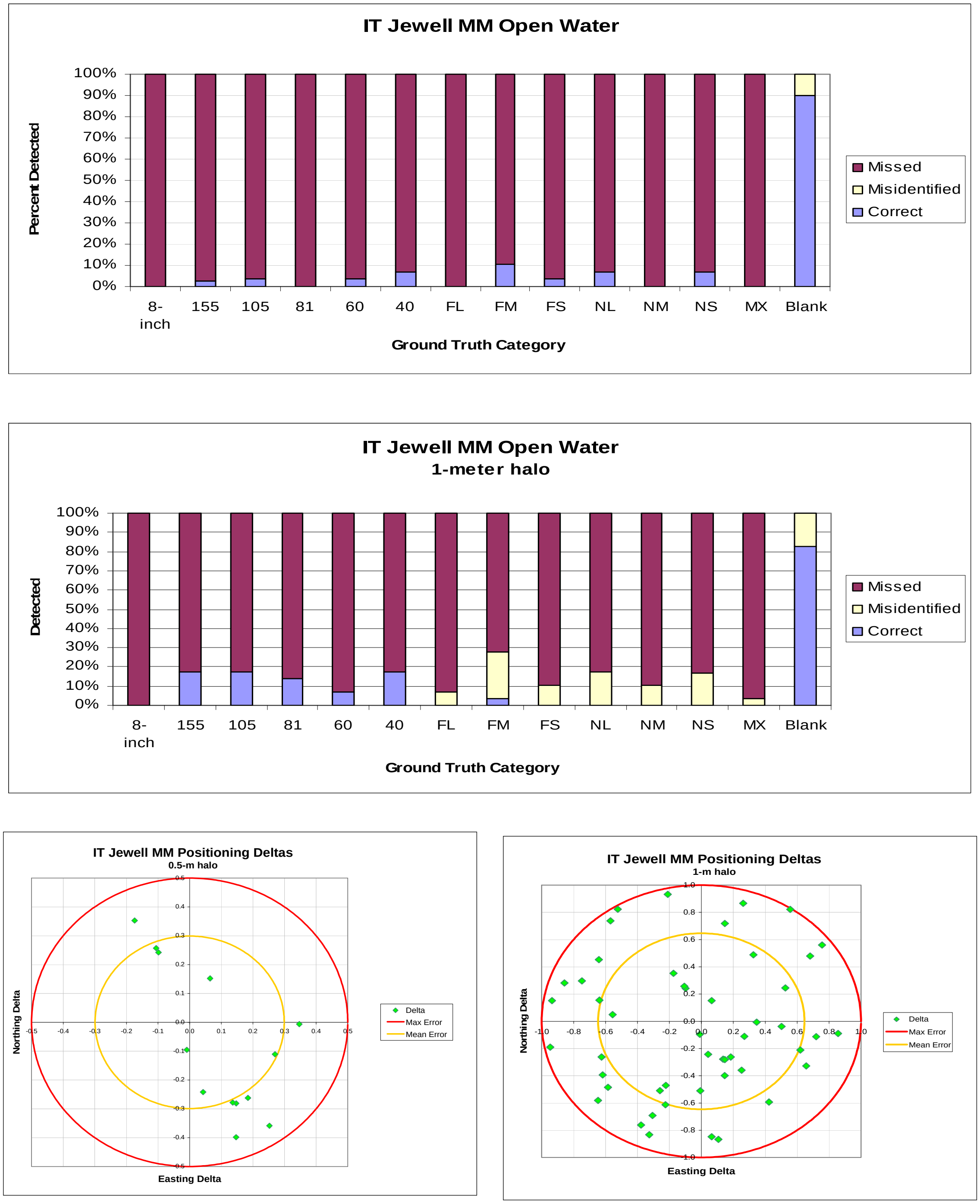


Figure 5-18. IT Jewell MM open water performance.

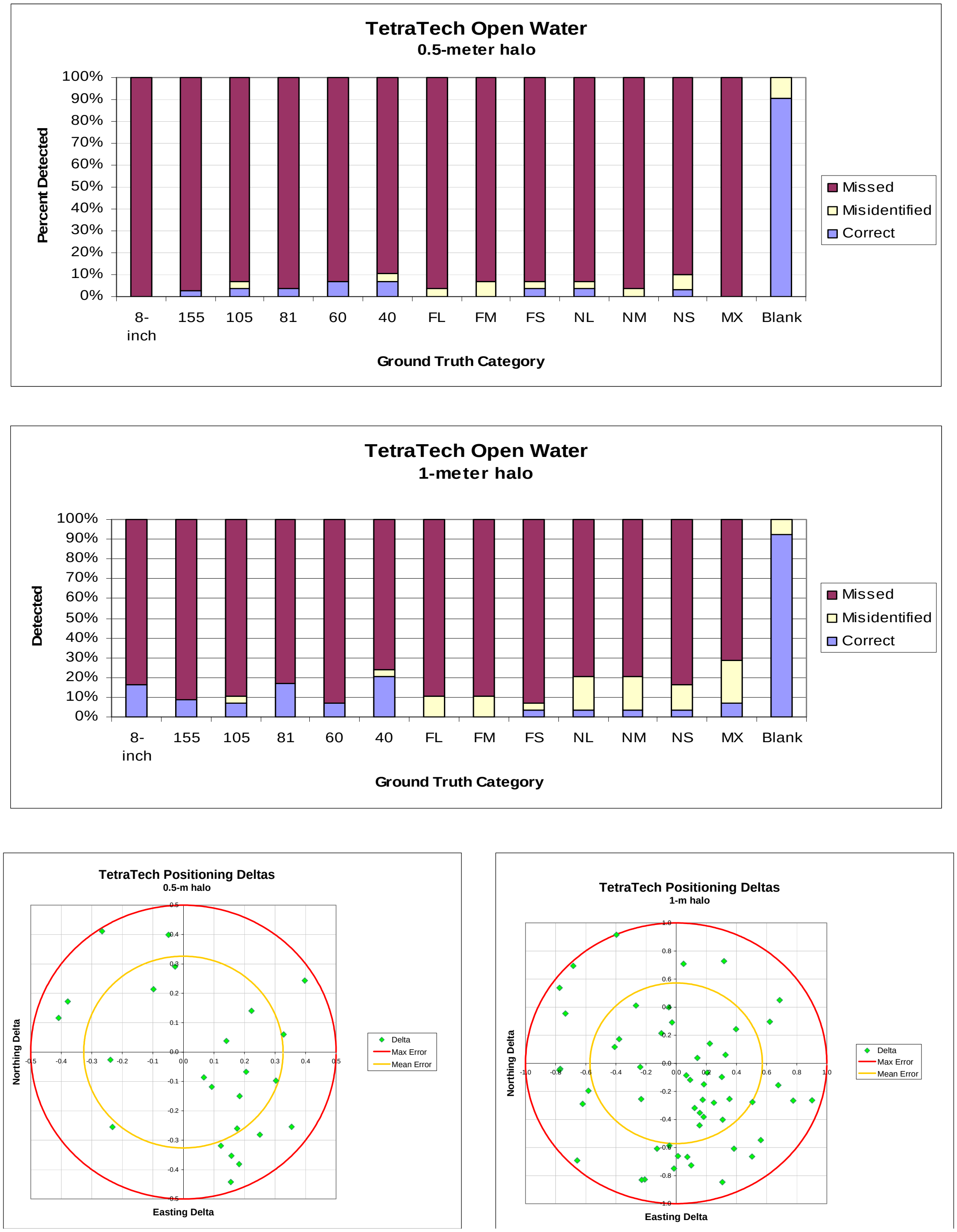


Figure 5-19. Tetra Tech open water performance.

5.3.5 Open Water Technical Assessment

As expected, the detection rate for all systems increased as the halo size increased from a 0.5- to a 1-meter radius. This suggests two possibilities: the occurrences of chance detections are higher in the larger surface area considered within the 1-meter halo, or there are location and/or positional errors in the systems.

5.3.5.1 True Detections

a. The probability of true detections was determined using the following analytical procedure[4]:

$$T = M - (U \cdot F \cdot \alpha)$$

where T = number of true matches:

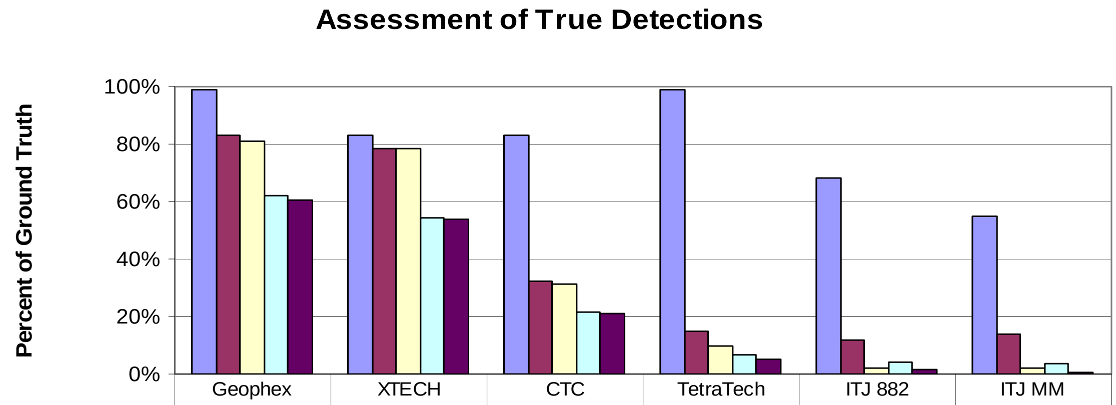
- U = true number of missed detections = *B* - *T*.
- B = number of items buried.
- F = true number of false detections = *D* - *T*.
- D = number of target declarations.
- α = (area of halo)/(area of site).
- M = sum of true + lucky matches.

This results in a quadratic in T whose solution is:

$$T = \frac{-\left(\frac{1}{\alpha} - B - D\right) + \sqrt{\left(\frac{1}{\alpha} - B - D\right)^2 - 4\left(BD - \frac{1}{\alpha}M\right)}}{2}$$

b. Knowing the area each system surveyed, the number of GT items that passed within the specified sensor swath, and the number of dig-list items that matched the GT allowed ATC to apply demonstrator-specific values to each variable. All declared GT items (ferrous and nonferrous clutter as well as ordnance) are used in this analysis to provide the most comprehensive assessment. This analysis was applied using both a 0.5- and a 1.0-meter halo. The results are shown in Figure 5-20.

[4] Andrews, A., V. George, T. Altshuler, and M. Mulqueen, "Results of the Countermine Task Force Mine Detection Demonstration at Fort A.P. Hill, Virginia, March 18-22, 1996," IDA paper P-3192, July 1996



	Geophex	XTECH	CTC	TetraTech	ITJ 882	ITJ MM
Targets Surveyed	99%	83%	83%	99%	68%	55%
1.0-M Matched	83%	78%	33%	15%	12%	14%
1.0-M True	81%	78%	31%	10%	2%	2%
0.5-M Matched	62%	54%	21%	7%	4%	4%
0.5-M True	61%	54%	21%	5%	1%	1%

Figure 5-20. Assessment of true detections.

c. Geophex, NAEVA/XTECH, and CTC/Foerster have the highest percentages of overall detections and the lowest percentages of chance detections. The opposite is true of the Tetra Tech and two IT Jewell systems.

5.3.5.2 Positional Errors

a. The data used in the 1-meter scatter graphs are plotted as histograms to show the location error distribution for each surveyor (fig. 5-21). Note that the scale on the ordinate axis varies between most demonstrators.

b. Target placement note: The pond was drained to allow the GT items to be emplaced and surveyed. Placement accuracy is estimated at ± 5 cm.

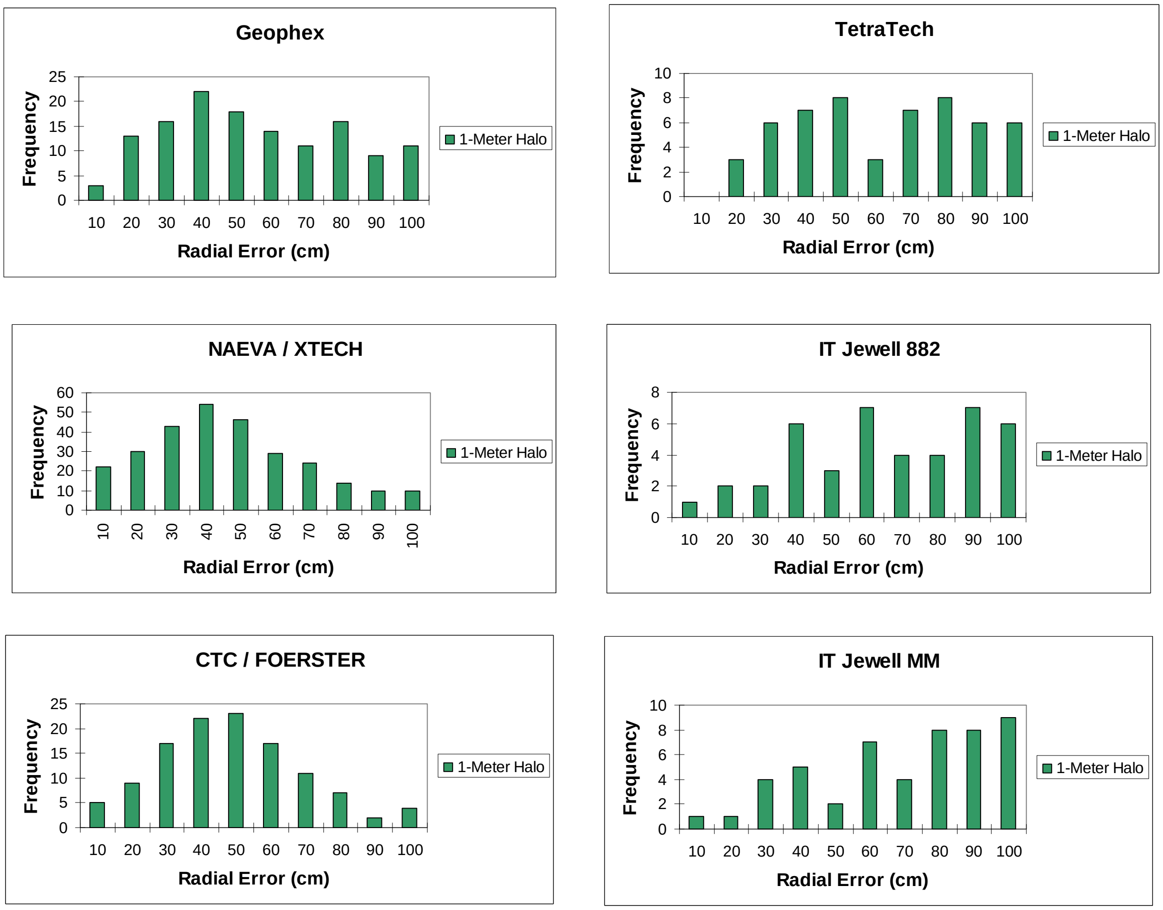


Figure 5-21. Radial error histograms.

c. The radial error histograms for Geophex, NAEVA/XTECH, and CTC show a right skewed location error distribution.

d. The low levels of true detections, the higher probability of chance detections, and the random reported item positions indicate that the Tetra Tech and the two IT Jewell systems cannot detect metal objects in this shallow water environment.

5.4 LITTORAL ZONE

5.4.1 Description

a. This is a sloping area on one side of the pond with vegetation growing into the water line. It contains several navigational and detection challenges. The GT items are underwater depths that range from 1 to 6 feet (0.3 to 1.8 m).

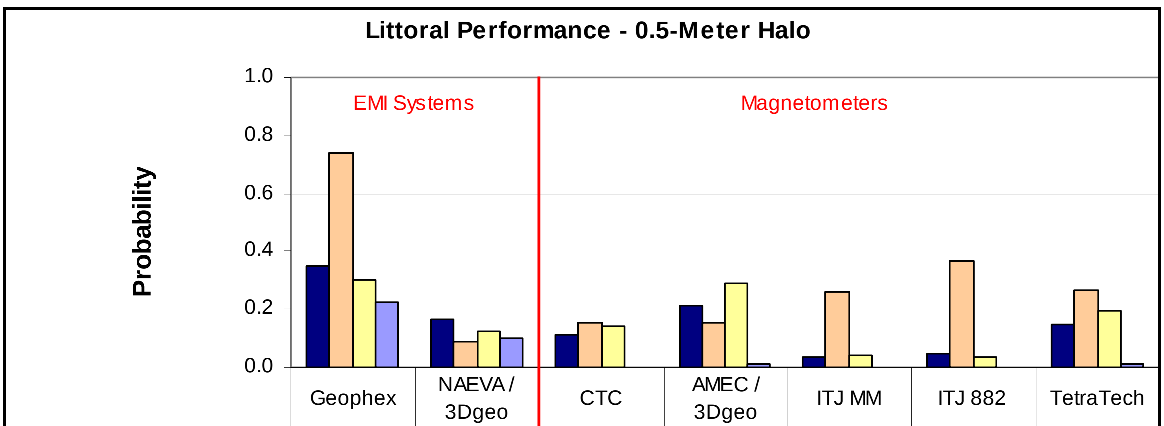
b. The GT in this area contains, as a subset, all of the items found in the blind grid. The GT burial depths in this area have also been randomized.

5.4.2 Purpose

Shorelines are a high-risk area for remediation efforts because many recreational activities take place where the water meets land. This challenge area provides an indication of the detection and discrimination abilities of each system in this type of area.

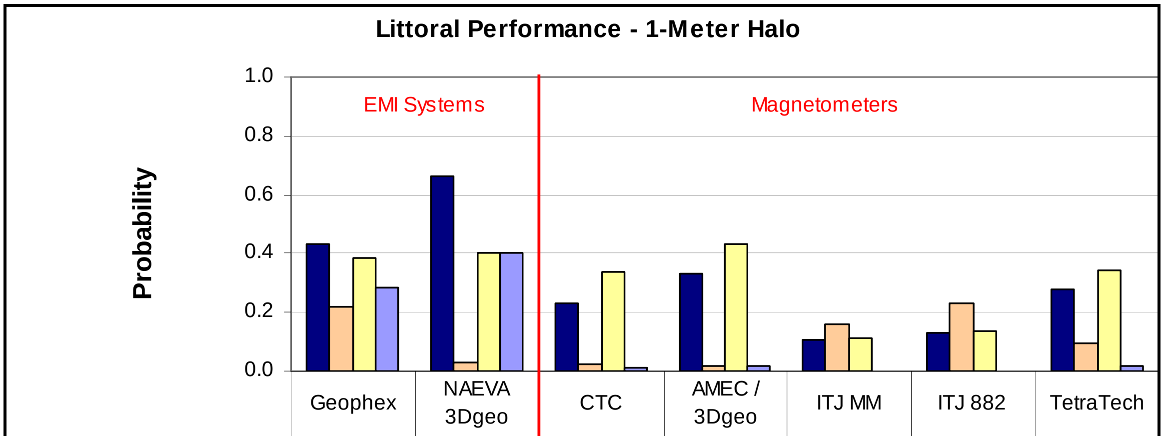
5.4.3 Response and Discrimination Findings

a. The performance results of the systems in this area are calculated using both a 0.5- and 1-meter detection halo. Summaries of the performance of the various demonstrators against the GT based on the terms and definitions in section 4 are shown in Figures 5-22 and 5-23.



	Geophex	NAEVA / 3Dgeo	CTC	AMEC / 3Dgeo	ITJ MM	ITJ 882	TetraTech
Metal Detected	0.35	0.17	0.11	0.21	0.03	0.05	0.15
False Metal Detection	0.74	0.09	0.16	0.15	0.26	0.37	0.27
Ord Correctly Identified	0.30	0.12	0.14	0.29	0.04	0.03	0.19
Clutter Correctly Identified	0.22	0.10	0.00	0.01	0.00	0.00	0.01

Figure 5-22. Littoral zone surveyor performance summary, 0.5-meter halo.



	Geophex	NAEVA 3Dgeo	CTC	AMEC / 3Dgeo	ITJ MM	ITJ 882	TetraTech
Metal Detected	0.43	0.66	0.23	0.33	0.11	0.13	0.28
False Metal Detection	0.22	0.03	0.02	0.02	0.16	0.23	0.10
Ord Correctly Identified	0.39	0.40	0.34	0.43	0.11	0.14	0.34
Clutter Correctly Identified	0.28	0.40	0.01	0.02	0.00	0.00	0.02

Figure 5-23. Littoral zone surveyor performance summary, 1-meter halo.

b. The preceding graphs show that most of the magnetometers have a higher probability of correctly identifying ordnance than detecting metal. All of the ordnance contains ferrous material whereas the metal category includes both ferrous and nonferrous items.

5.4.4 Detailed Findings

a. The detailed detection and discrimination results shown in Figures 5-24 through 5-30 are based on the procedures in section 4. The scatter graphs show the difference between the emplaced GT center and the surveyor-reported position for all true metal detections (ordnance and clutter). The positioning deltas are calculated using the square root of the sum of the squares of the northing and easting errors.

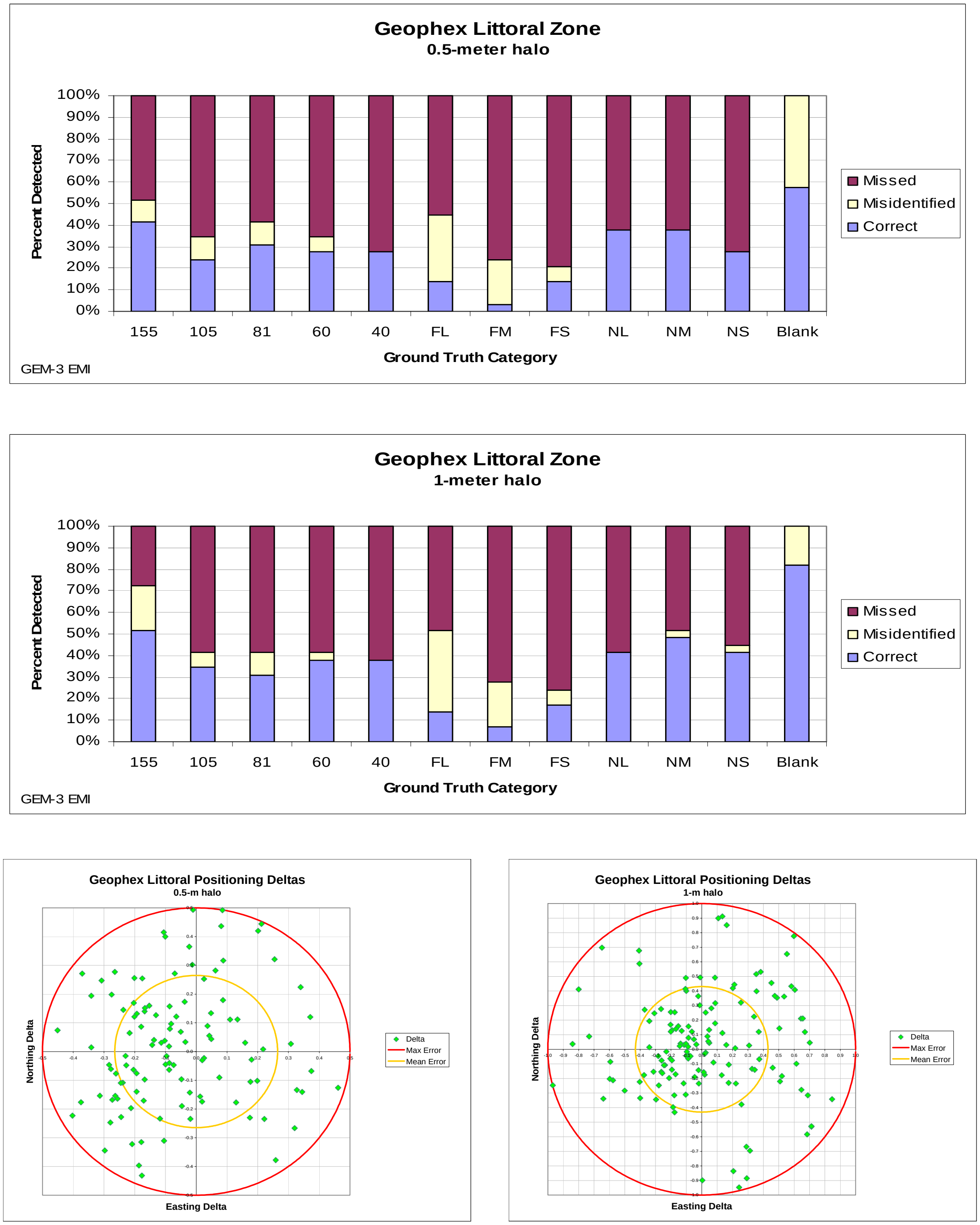
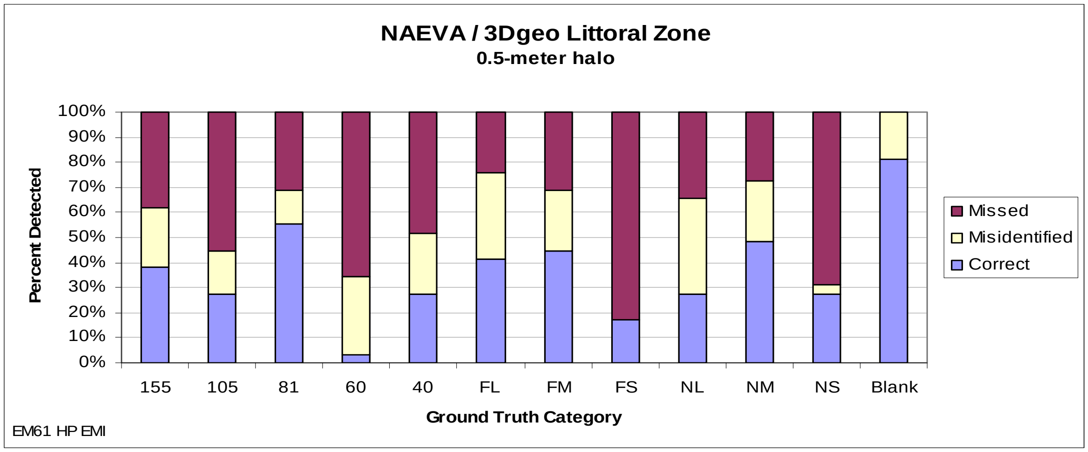
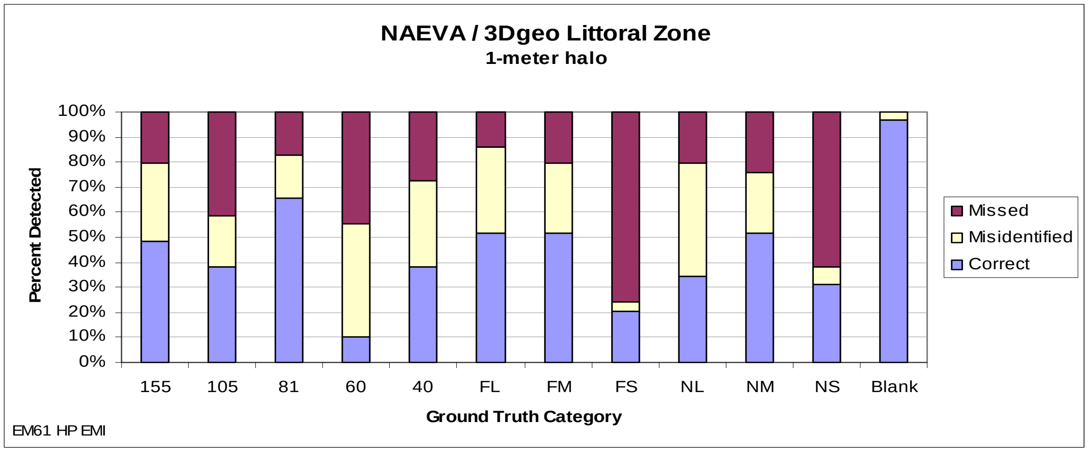
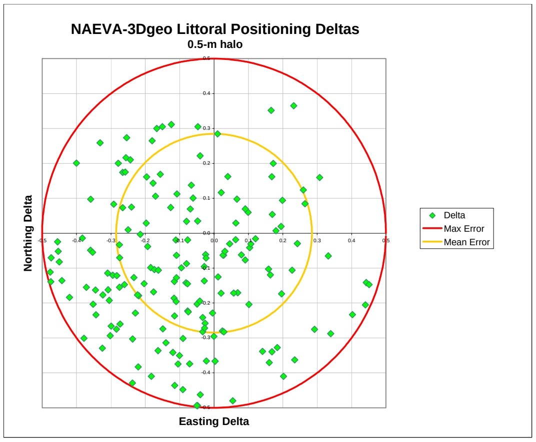


Figure 5-24. Geophex littoral zone performance.

Note: Geophex used the same sensor array mounted on a different survey platform in the littoral zone (para 2.1.2).







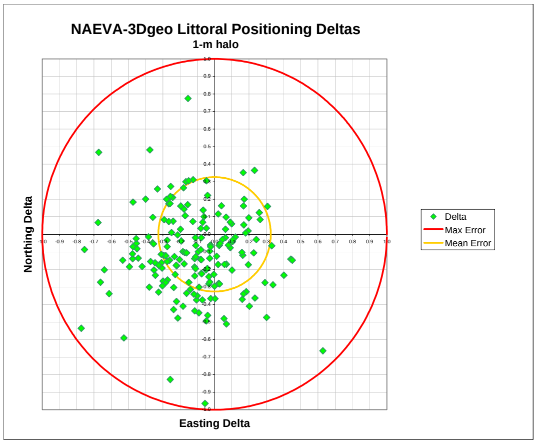


Figure 5-25. NAEVA/3Dgeo littoral zone performance.

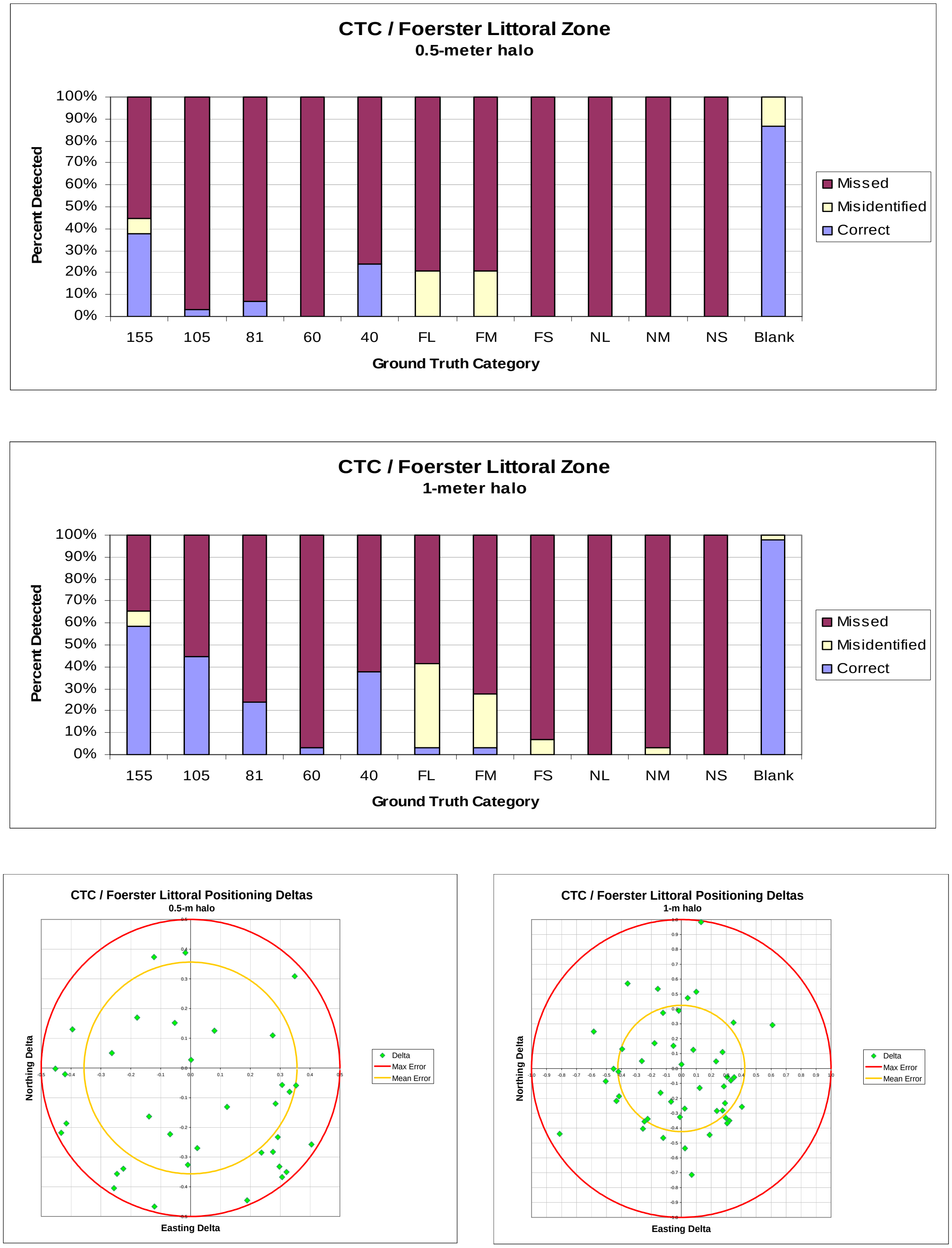
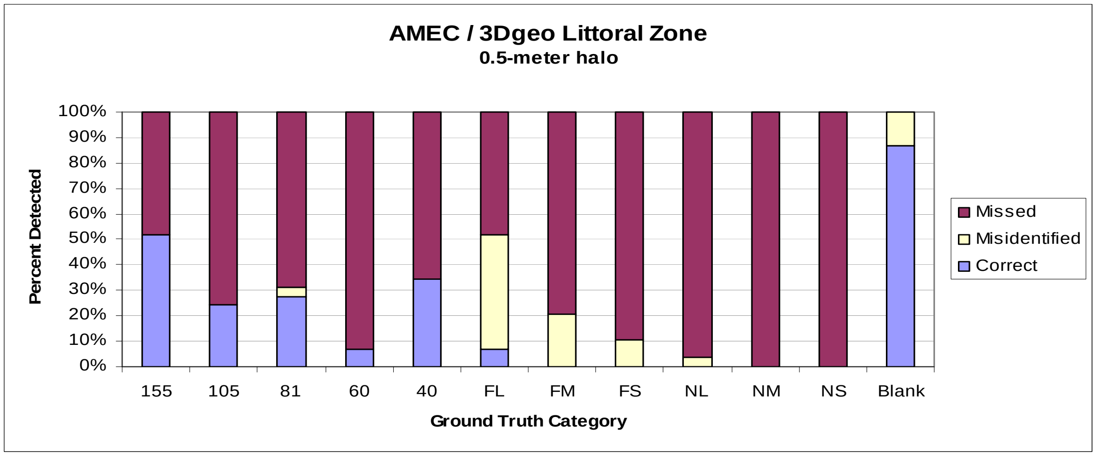
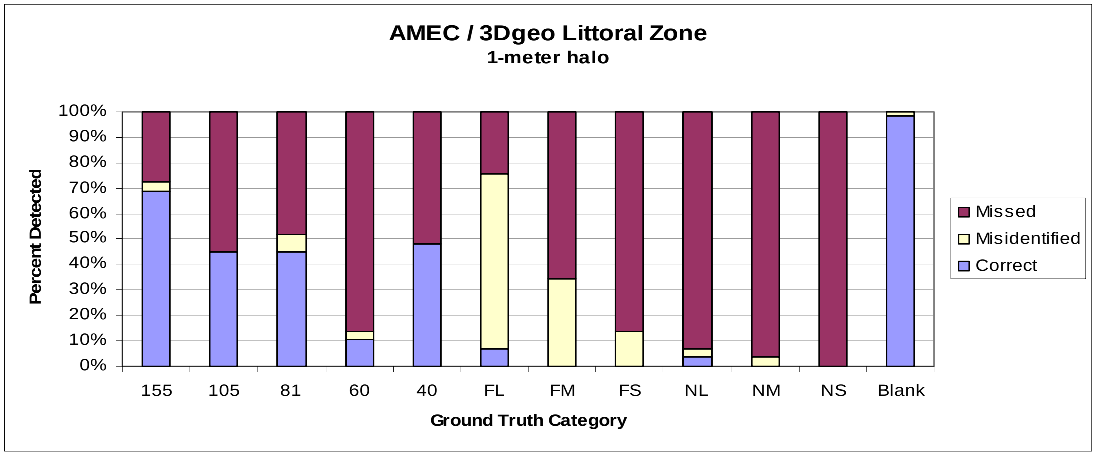
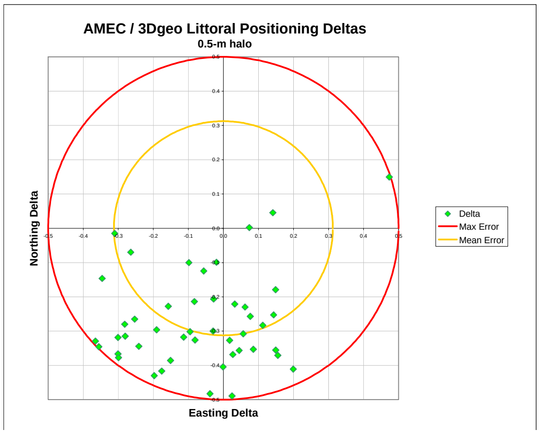


Figure 5-26. CTC/Foerster littoral zone performance.







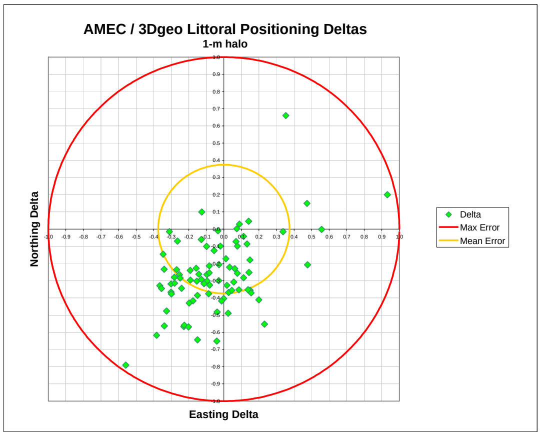


Figure 5-27. AMEC/3Dgeo littoral zone performance.

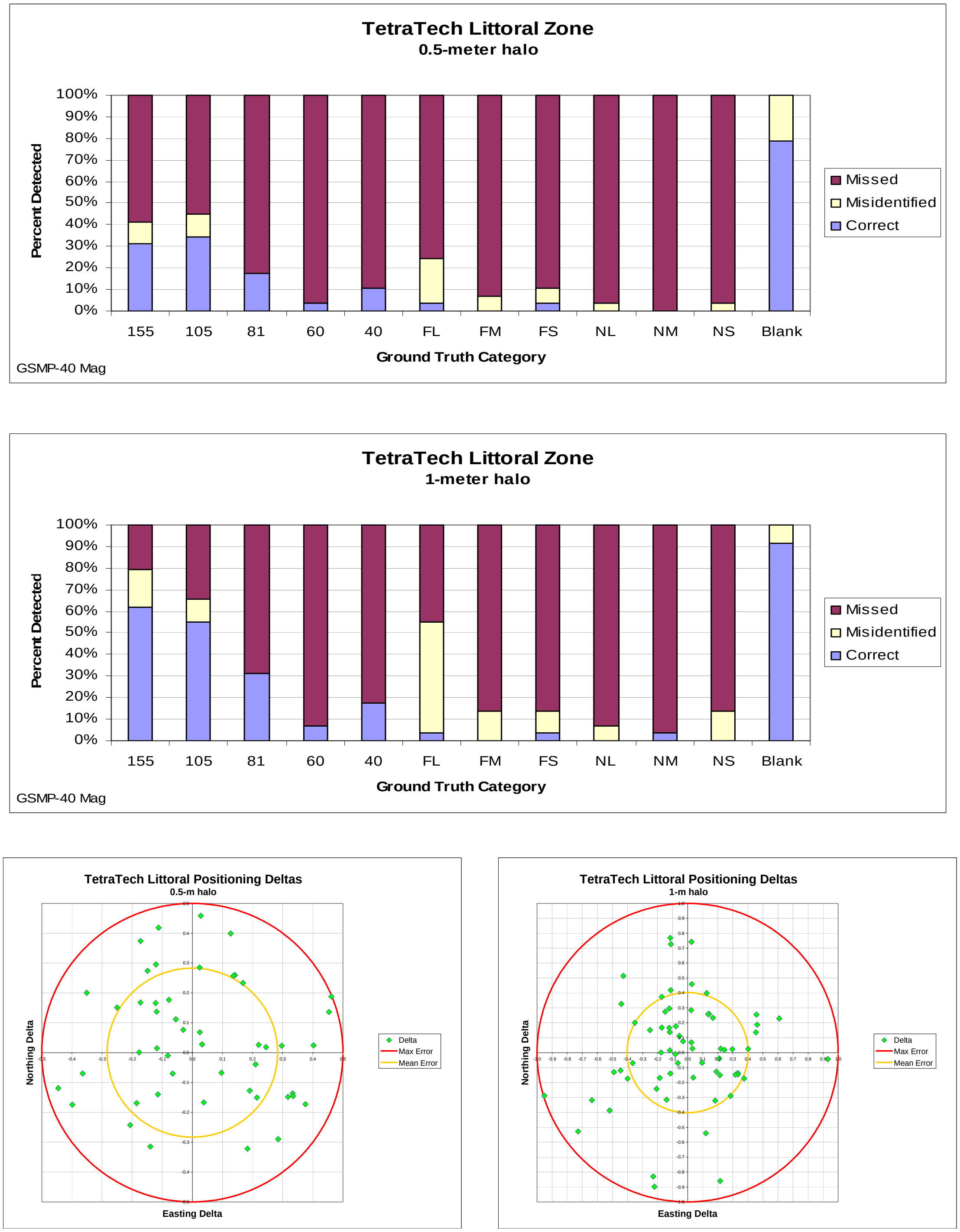


Figure 5-28. Tetra Tech littoral zone performance.

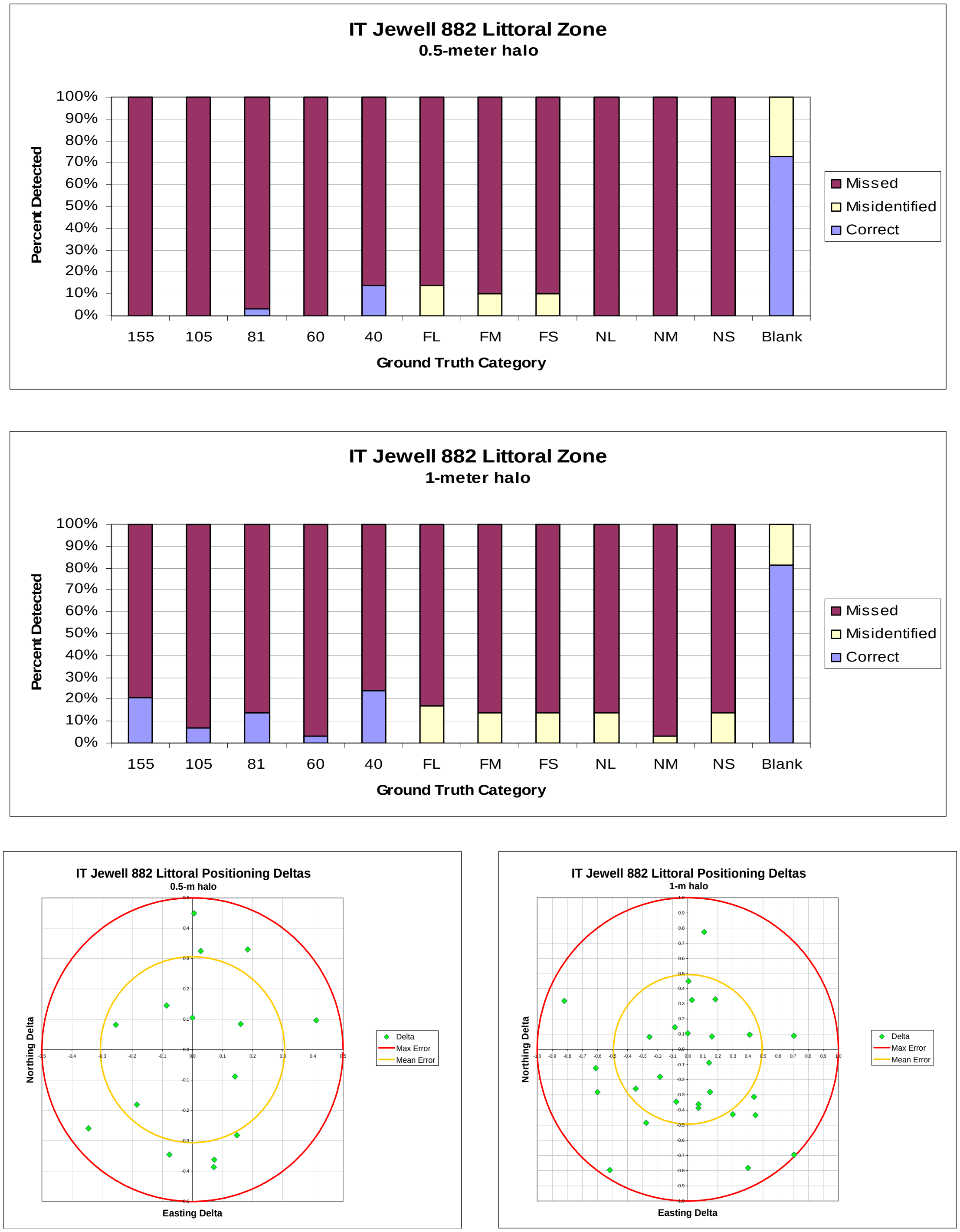


Figure 5-29. IT Jewell 882 littoral zone performance.

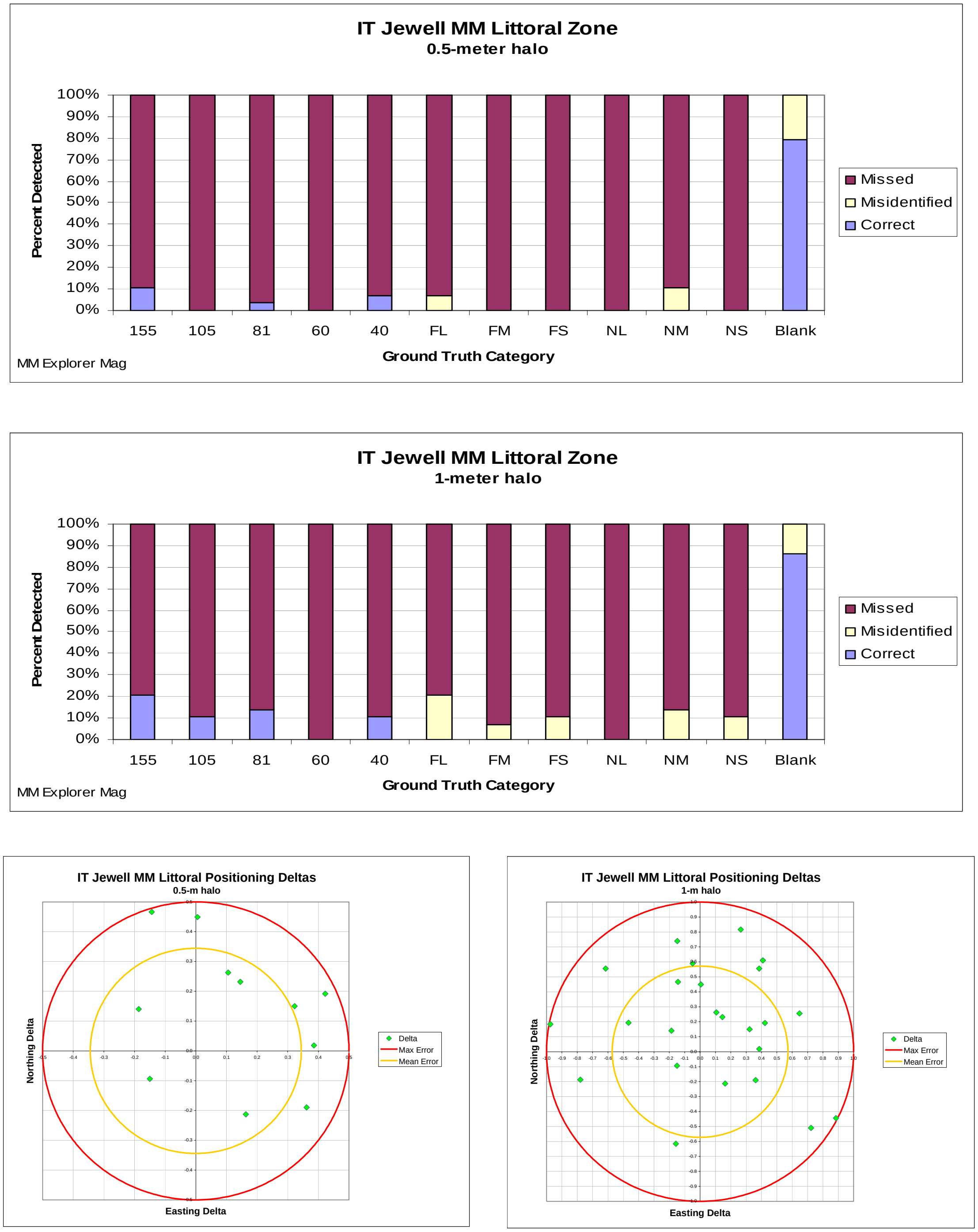
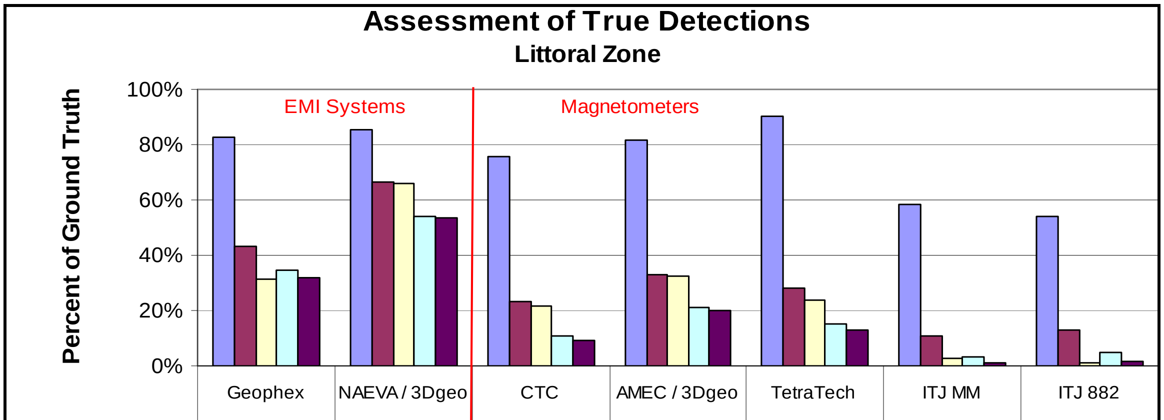


Figure 5-30. IT Jewell MM littoral zone performance.

5.4.5 Littoral Zone Technical Assessment

5.4.5.1 True Detections

a. The probability of chance detections was analyzed using the method described in paragraph 5.3.5.1. The results are in Figure 5-31.



	Geophex	NAEVA / 3Dgeo	CTC	AMEC / 3Dgeo	TetraTech	ITJ MM	ITJ 882
Items Surveyed	83%	85%	76%	82%	90%	58%	54%
1.0-M Matched	43%	66%	23%	33%	28%	11%	13%
1.0-M True	31%	66%	22%	33%	24%	3%	1%
0.5-M Matched	35%	54%	11%	21%	15%	3%	5%
0.5-M True	32%	53%	9%	20%	13%	1%	2%

Figure 5-31. Assessment of true detections.

b. The Geophex system has the highest difference between matched and true detections of all the systems using a 1-meter halo.

c. The shallower water in this area placed Tetra Tech's magnetometers closer to the GT targets. This resulted in a higher detection rate for this system when compared with its performance in the other test areas.

5.4.5.2 Positional Errors

a. The data used in the 1-meter scatter graphs were replotted as histograms to show the location error distribution for each surveyor (fig. 5-32). Note that the scale on the ordinate axis varies between most demonstrators.

b. GT placement error is estimated at ± 5 cm.

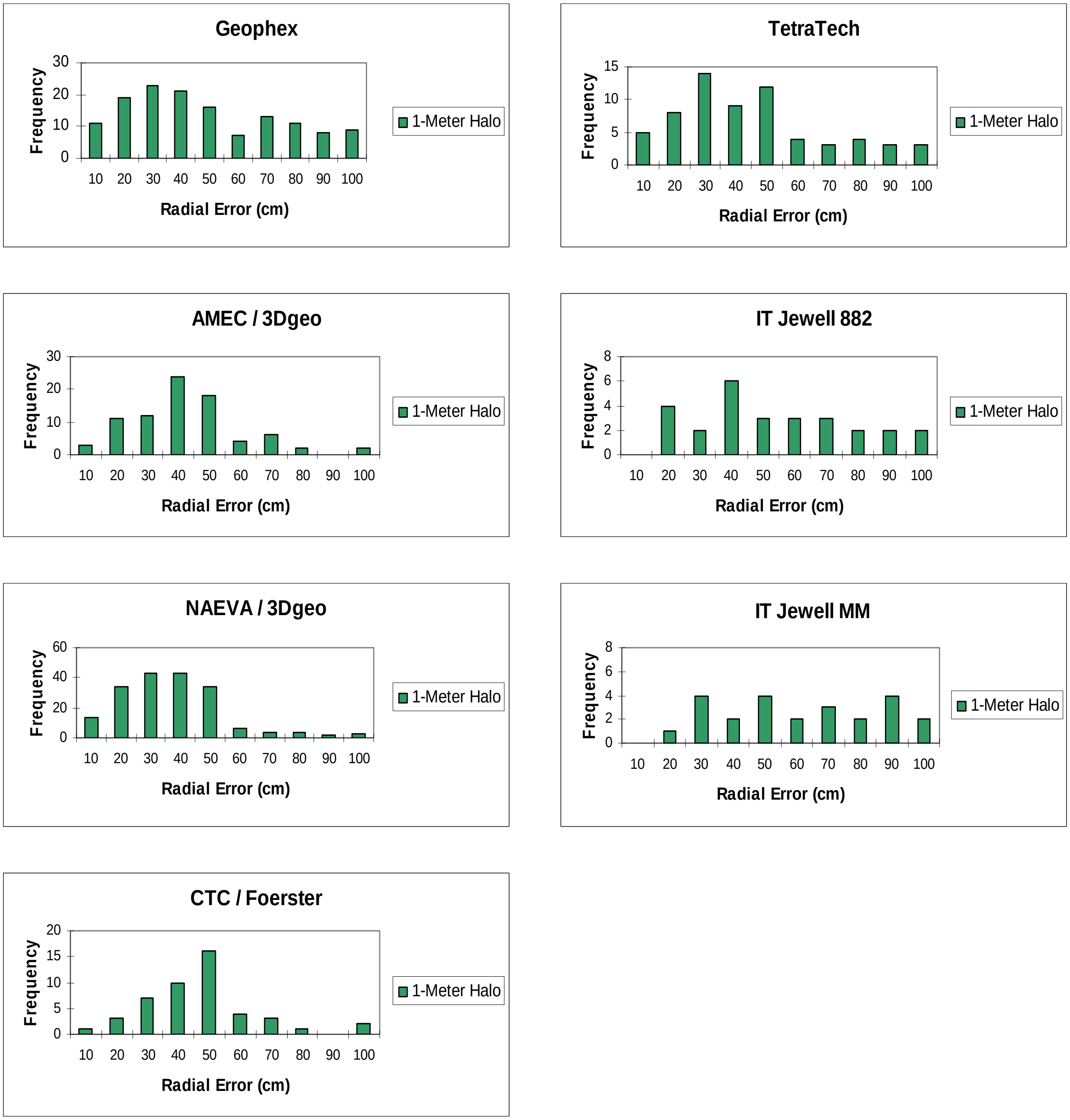


Figure 5-32. Littoral zone radial error histograms.

c. The location error histograms for most systems tested in this area have a right skewed distribution. The IT Jewell histograms do not show a normal error distribution.

d. Comparing the littoral zone and open water histograms shows the following:

(1) The CTC/Foerster and Geophex systems show a similar error distribution pattern.

(2) The Tetra Tech distribution shifted to the left based on the improved detection results.

(3) The IT Jewell systems still showed a random distribution.

(4) The 3Dgeo platforms did not survey the open water area.

5.5 DEEPER WATER AREA

5.5.1 Description

During the construction of this site, the pond was drained, existing metallic items were removed, and then the GT items were emplaced. The location now called the deeper water area represents a low-lying area that did not dry out enough to survey and clear. A group of 155-mm projectiles was distributed across the ground surface in this area. The water depth varies from 10 to 14 feet (3.1 to 4.3 m).

5.5.2 Purpose

This area provides information on the probability of detecting large caliber ordnance in deeper water.

5.5.3 Response and Discrimination Findings

a. The probability of detecting 155-mm projectiles by six different systems is summarized in Figure 5-33. Damage to the deeper water site prevented two of the demonstrators (NAEVA/3Dgeo and AMEC/3Dgeo) from surveying this area.

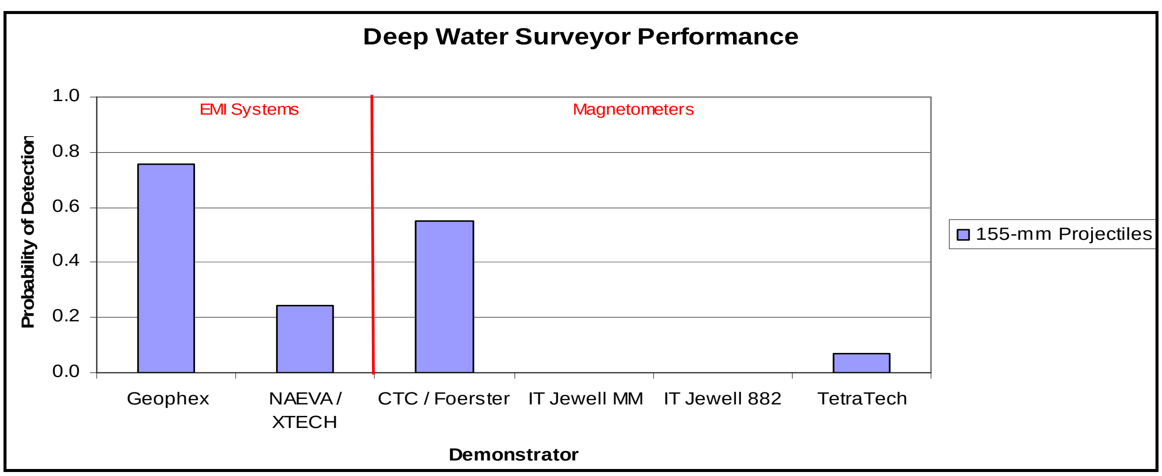


Figure 5-33. Probability of detection in the deeper water area.

b. Both EMI systems are on bottom-riding sleds. Geophex surveyed 98 percent of the test area; NAEVA/XTECH surveyed 64 percent of the area. On the magnetometer side, CTC lowered their probe depth to 9 feet (2.7 m) and surveyed 94 percent of the test area; the IT Jewell systems surveyed 70 and 64 percent, respectively, and Tetra Tech covered 98 percent of the test area.

5.5.4 Detailed Findings

None.

5.5.5 Deeper Water Technical Assessment

The data obtained in this area further demonstrate the effect of sensor to target distance in the detection process.

a. The probability of detection for the surface and subsurface systems is less than 0.1 and 0, respectively.

b. The CTC/Foerster system achieved its highest probability of detection in this area. Two factors contributed to this success: lowering the magnetometers to bring them closer to the ordnance items, and their discrimination process does well in identifying 155-mm projectiles.

c. Geophex, a bottom-riding system, obtained the highest probability of detection in this area.

d. The NAEVA/XTECH system had difficulty maneuvering in this area due to its towing method and the proximity of this area to the pond bank. The design of this bottom-riding platform dislodged several projectiles from their emplaced positions. This may have distorted the EM signatures and contributed to the low probability of detection. Chronologically, this was the last demonstrator to survey this area.

5.6 TECHNICAL ASSESSMENT OF ALL SYSTEMS ACROSS ALL TEST AREAS

a. The shallow water systems have been grouped by sensor operating principle, either EMI or magnetometer in this section. This is an oversimplification in that the detection and discrimination processes not only depend on the type of sensor but also on the design of the survey platform, the data collection and reduction process, QA/QC measures, and many other factors. However, this provides a structure for this assessment.

b. The two systems demonstrated by IT Jewell did not perform well in any of the test areas. The high probability of false detections (fig. 5-3, 5-12, 5-13, 5-22, and 5-23), the low probability of true detections (fig. 5-20 and 5-31), and the distribution of the measurement error in the histograms shown as part of Figures 5-21 and 5-32 indicate that these two systems cannot detect UXO in a shallow water environment.

c. The system provided by Tetra Tech was intended to be a dual-mode system, combining both acoustic and magnetic geophysical technologies. Problems reducing the sonar data required this demonstrator to develop the dig list based solely on magnetometer data. This magnetometer system cannot detect UXO in water depths greater than 8 feet (2.4 m) based on most of the same figures referenced for the IT Jewell systems. The shallower water (≥ 6 feet (1.8 m)) in the littoral zone placed the magnetometers closer to the GT targets. This reduced standoff distance produced more reliable detection results (fig. 5-23, 5-31, and 5-32).

d. The two IT Jewell systems are not included in the remainder of this assessment. Only the data from the Tetra Tech littoral zone are assessed further.

e. The rest of this assessment consolidates the performance of the remaining systems across test areas in terms of the probability of metal detection and the probability of discriminating between ordnance and clutter. These values are obtained using the procedures discussed in section 4. This information is summarized in Figures 5-34 and 5-35..

f. Not all test areas were surveyed by all systems.

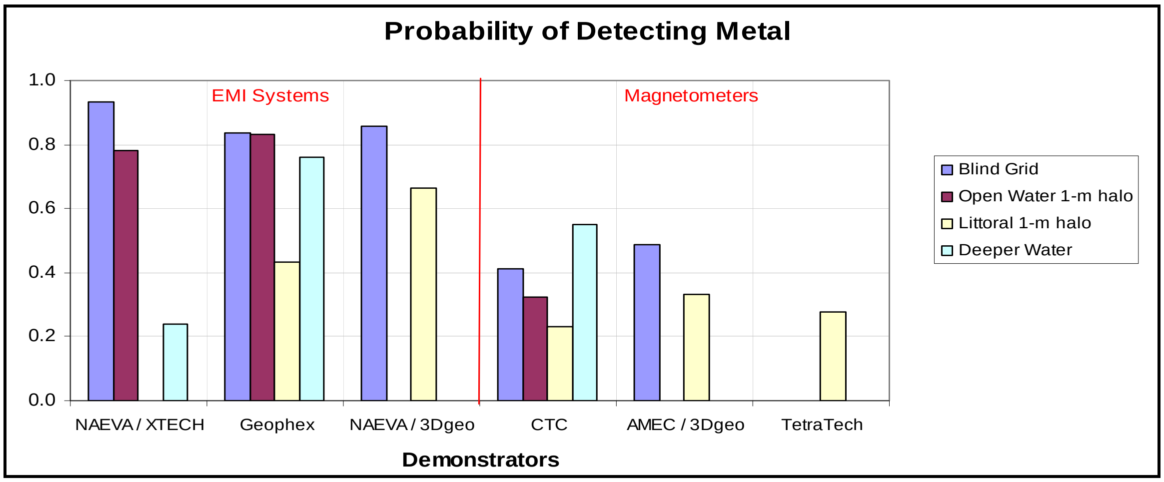


Figure 5-34. Demonstrator's probability of detecting metal by test areas.

g. The EMI systems detected more objects that are metallic in all of the test areas compared with the magnetometer results. The detailed detection and discrimination charts show that there are two contributing factors, the detection of a greater number of nonferrous items and the detection of a greater number of the medium and small ferrous clutter items.

h. The probability of metal detection for all systems decreases as the survey platforms move from the blind grid to the open water and then to the littoral zone. There are two identified causes. The first is the ability of the systems to get into an area to survey it (table 3-1). The second is the positioning errors shown in Figures 5-21 and 5-32.

i. The GT targets in the deeper water area consisted of 155-mm projectiles only. Limiting the detection requirements to this large projectile accounts for the improved performance in this area. The NAEVA/XTECH system only surveyed 65 percent of this test because of maneuverability limitations (para 2.5.5).

j. The probabilities of correctly discriminating found objects as either ordnance or clutter are shown in Figure 5-35.

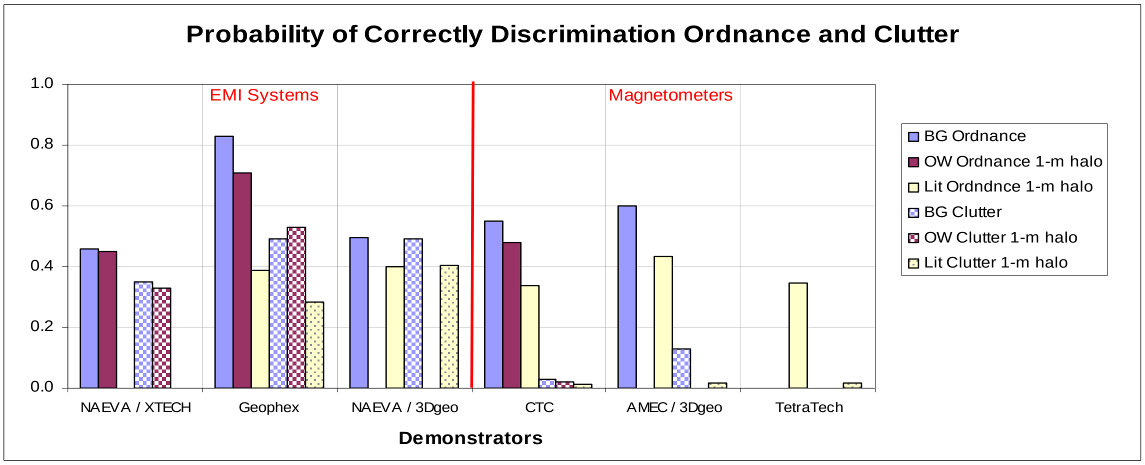


Figure 5-35. Demonstrator's probability of correctly discriminating ordnance from clutter.

k. In the three main test areas (deeper water excluded), the EMI systems always performed better than the magnetometer systems in terms of detecting GT items. In terms of correctly discriminating clutter items, the EMI systems again showed better results, but then, they found more of the clutter items than the magnetometer systems. These trends do not hold in the ordnance discrimination process. Only one of the three EMI systems has a higher probability of identifying ordnance than the magnetometer systems in the blind grid area. The EMI systems correctly identified more ordnance items in the open water test area, whereas a magnetometer system had the highest probability of correctly identifying ordnance in the littoral zone.

l. Three conclusions can be drawn from this summary and a review of the detailed detection and discrimination plots:

(1) If your objective is to detect ferrous ordnance items, there is not much difference between the bottom-riding magnetometers and EMI systems.

(2) The current state of the discrimination process lags far behind the detection process.

(3) The more objects you find, the better your discrimination process needs to be.

SECTION 6. SITE REMEDIATION COMPARISON

This section provides a comparative analysis of demonstrator performance in terms of the two main components of site remediation: survey cost and the amount of UXO removed.

6.1 SURVEY COSTS

a. A standardized system was developed to compare on-site costs between demonstrators. This system recorded the amount of time spent on equipment setup and calibration, site survey and any resurvey time, downtime due to system malfunctions and maintenance requirements, and demobilization. The site survey time included daily setup/stop time, collecting data, breaks/lunch, downtime for equipment/data checks or maintenance, downtime due to failure, and downtime due to weather. Survey personnel were placed in one of three standardized labor categories, each with a standardized rate: supervisor ($95.00/hr), data analyst ($57.00/hr), and site support ($28.50/hr). The total survey costs equal the number of personnel, by labor category, multiplied by the number of on-site hours.

b. The process described above provides an on-site cost guide to compare the performance of all vendors at the shallow water site. It is not a true indicator of survey costs. Many other expenses have not been included, such as travel costs, per diem, off-site data processing and analysis, company administrative overhead, and profit.

c. Calculating the area surveyed by each demonstrator is described in section 3. The total area surveyed by each demonstrator is shown in Table 6-1. To determine the number of acres surveyed per day, the total number of hours spent at the test site was divided by 8 (converts to 8-hr days). The number of acres was then divided by the number of 8-hour days. The cost per acre was determined by dividing total survey cost by the same number of acres.

TABLE 6-1. SURVEY COSTS

	Geophex	**Tetra Tech**	**CTC Foerster**	**NAEVA XTECH**	**NAEVA 3Dgeo**	**AMEC 3Dgeo**	**IT Jewell**
Calculated survey cost[a]	$11,335	$9,748	$7,707	$12,489	$5,998	$6,207	$5,884
Area surveyed (acre[b])	5.35	5.57	4.25	3.7	[c]2.8	[c]2.8	[d]7.34
Time on-site (8-hr days)	8.9	6.7	5.2	7.5	4.15	4.15	4.43
Acres per day	0.60	0.83	0.82	0.49	0.67	0.67	1.66
Cost per acre	$2,119	$1,750	$1,813	$3,375	$2,142	$2,217	$802

[a]U. S. dollars
[b]1 acre = 4047 square meters
[c]Only the blind grid and littoral areas were available for surveying.
[d]Site surveyed twice using different sensors.

d. Detailed information on the time spent on-site and calculated survey costs is in the individual demonstration reports.

6.2 VIRTUAL SITE REMEDIATION

a. Each demonstrator provided a dig list for the areas they surveyed. This list contained the coordinates and discriminated classification (either ordnance or clutter) of the items found. Comparing the ordnance and clutter declarations from these lists with the emplaced GT items in each test area permits a virtual site remediation analysis.

b. The results of that virtual recovery operation in terms of the percentage of ordnance and clutter recovered, or left behind, and the percentage of empty digs (empty ordnance declarations) are shown in Figures 6-1 through 6-6. Two charts are shown for each area. The first is based on sensor response information, which assumes that everything detected will be recovered. The second is based on the discrimination process, which assumes that only items declared ordnance would be recovered. All graphs use the entire range of GT targets in each test area.

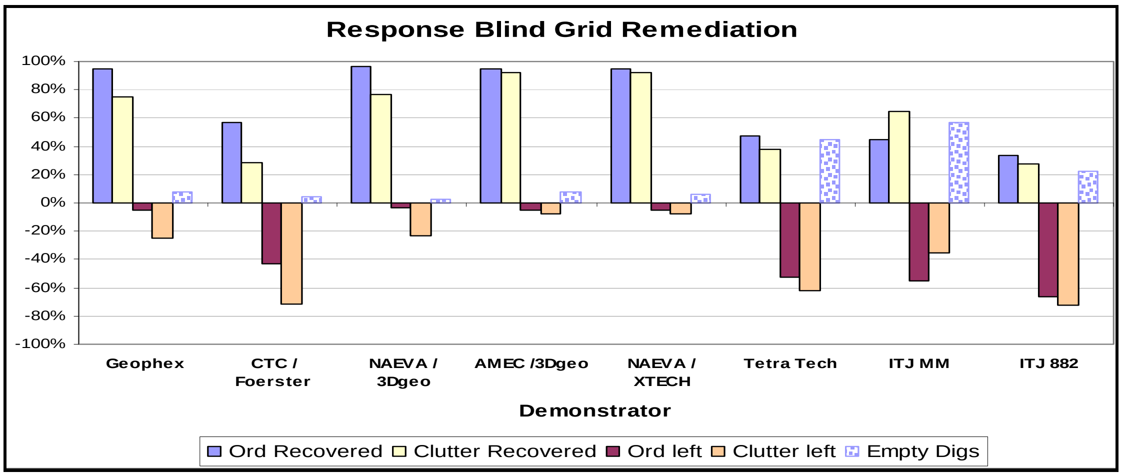


Figure 6-1. Response blind grid remediation.

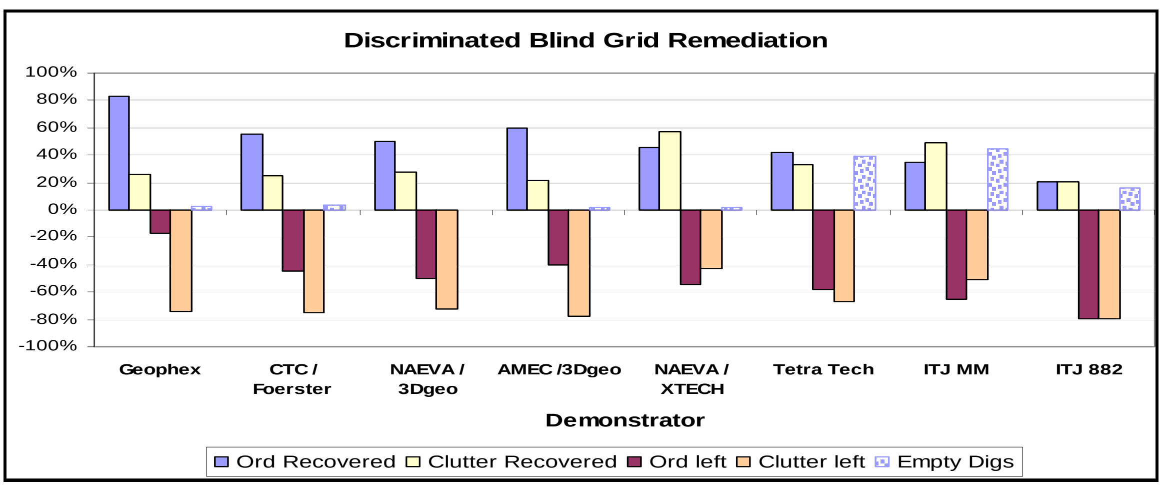


Figure 6-2. Discriminated blind grid remediation.

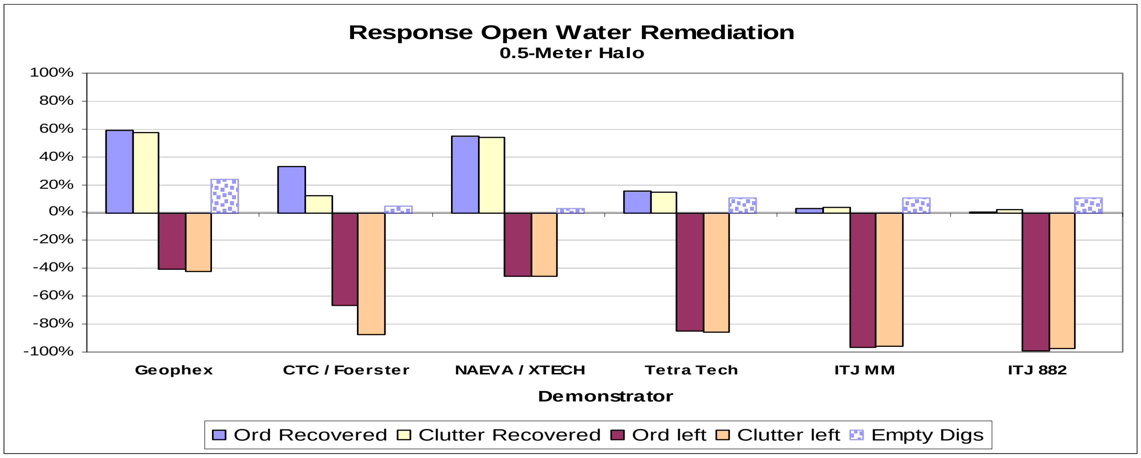


Figure 6-3. Response open water (0.5-m halo) remediation.

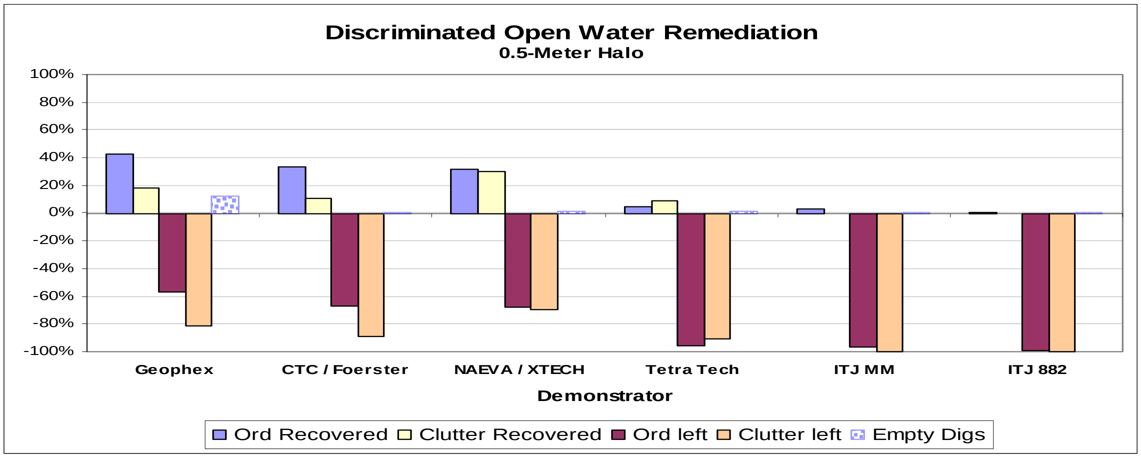


Figure 6-4. Discriminated open water (0.5-m halo) remediation.

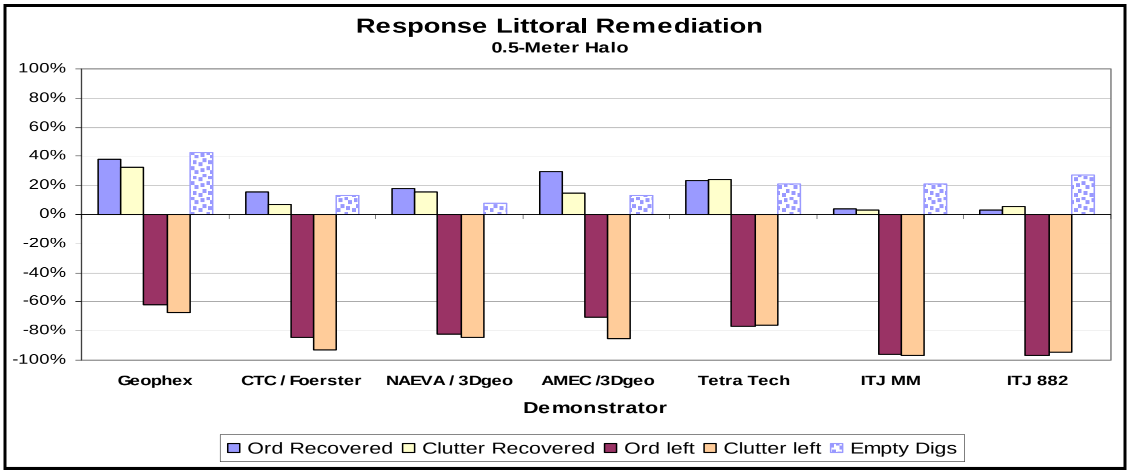


Figure 6-5. Response littoral remediation.

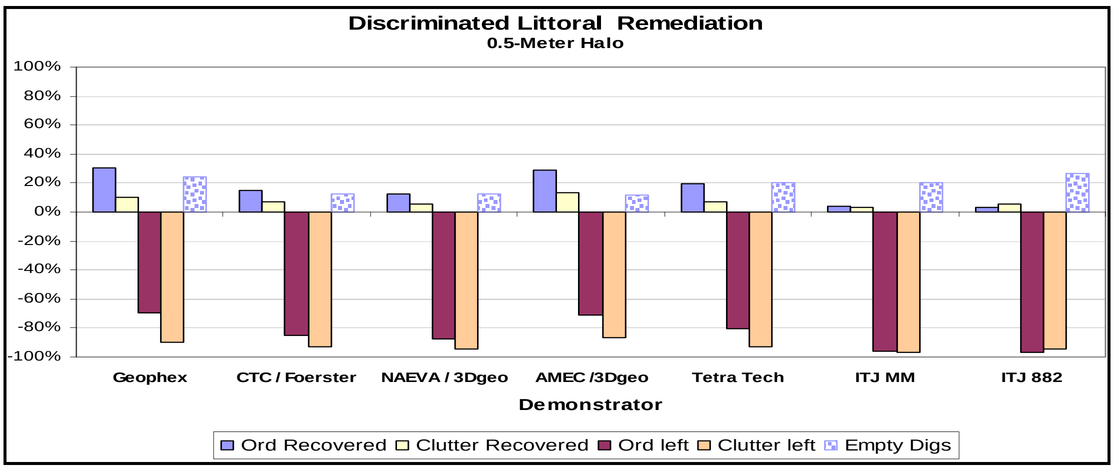


Figure 6-6. Discriminated littoral remediation.

6.3 TECHNICAL ASSESSMENT

a. The survey costs shown in Table 6-1 do not reflect the contract amounts allocated to each demonstrator. Striking a balance between proprietary cost information and accurate reporting, the true cost of each survey is approximately an order of magnitude higher than the values shown.

b. The response blind grid remediation chart shows that four of the systems tested are capable of locating more than 90 percent of the ordnance in this test area. Applying the discrimination routines that differentiate ordnance from clutter reduces the reported percentage of ordnance items from 12 to 50 percent, depending on the surveyor. On the other side of the discrimination routine, the percentage of clutter items that would be recovered is reduced from 35 to 70 percent, again depending on the surveyor. Recovering all of the metal objects reported by a surveyor reduces the risk but increases the costs; recovering only the discriminated ordnance items reduce the cost and apparent risk. Improvements in the discrimination routines will benefit both sides of this equation.

c. Response and discrimination recovery percentages decrease as the systems move from the blind grid to the open water and littoral areas.

SECTION 7. APPENDIXES

APPENDIX A. USAEC TERMS, DEFINITIONS AND SCORING PROCEDURES

The terms, definitions and scoring procedures in this appendix generated the results shown in Appendix B.

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. 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 projected length of the ordnance onto the ground plane plus 1 meter.

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 selects the 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 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: An anomaly in a blind grid cell that contains neither emplaced ordnance nor an emplaced clutter item. An anomaly location in the open water or littoral 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 water 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: An anomaly in a blind grid cell that contains neither emplaced ordnance nor an emplaced clutter item. An anomaly location in the open water or littoral 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.[a] 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.

[a] Strictly speaking, ROC curves plot the P_d versus P_{ba} over a predetermined 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 water scenario, each system suppresses its signal strength reports until some bare-minimum signal response is received by the system. Consequently, the open water ROC curves do not have information from low signal-output locations, and, furthermore, different contractors report their signals over a different set of locations of the pond. 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.

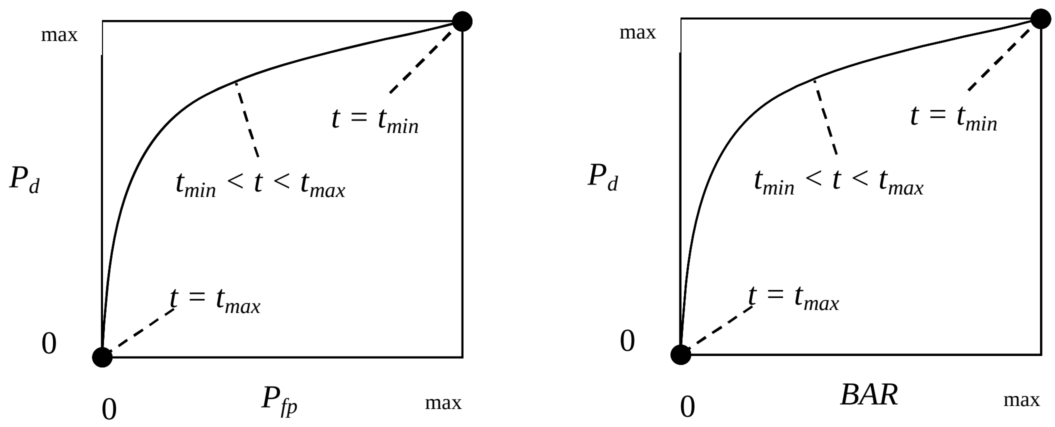


Figure A-1. ROC curves for open-site testing. Each curve applies to both the response and discrimination stages.

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 water: $R_{ba} = 1 - [BAR^{disc}(t^{disc})/BAR^{res}(t_{min}^{res})])$

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[5].

A one-sided 2 x 2 contingency table is used in the Shallow Water Site Program to compare each area (open water, littoral, deep water) to the blind grid since each area introduces a water feature that makes it potentially more difficult to survey than the Blind Grid.

The one-sided 2 x 2 contingency table is used 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 feature introduced. A two-sided 2 x 2 contingency table is used to compare performance between any two of the test sites other than the blind grid, to determine if there is reason to believe that the proportion of ordnance correctly detected/discriminated by demonstrator X's system is significantly different between those two test sites.

The test statistic of the 2 x 2 contingency table is the Chi-square distribution with one degree of freedom. For the one-sided test, a significance level of 0.05 is chosen which sets a critical decision limit of 3.84 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 Exact Test is used and the critical decision limit is the chosen significance level, which is 0.05 for one-sided tests and 0.10 for two-sided tests. With Fischer's test, if the test statistic (p-value) is less than the critical value, then the null hypothesis of similar performance is rejected in favor of the alternative hypothesis: significantly greater than for the one-sided case or significantly different for the two-sided case.

Shallow-water UXO Detection Test Site examples, where blind grid results are compared to those from the open water and littoral sites and the non-grid sites (open water and littoral) are compared to each other as follows. It should be noted that a significant result does not prove a cause and effect relationship exists between the change in survey area and sensor performance; however, it does serve as a tool to indicate that one data set reflects relatively degraded system performance of a large enough scale 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

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

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 areas using the same system (results indicate the number of ordnance detected divided by the number of ordnance emplaced):

	Blind grid	Open water	Littoral
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 WATER. 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 water. 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 (p-value) 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 water relative to results from the blind grid using the same system.

P_d^{disc}: BLIND GRID versus OPEN WATER. 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 out of 10 emplaced ordnance items were correctly discriminated as such in open water testing. Those four values are used in the Chi-square Contingency Test to calculate a test statistic of 1.12. Since the test statistic is less than the critical value of 3.84, the two discrimination stage detection rates are considered to be not significantly different at the 0.05 level of significance.

P_d^{res}: BLIND GRID versus LITTORAL. Using the example data above to compare probabilities of detection in the response stage, 100 out of 100 and 20 out of 33 are used to calculate a test statistic (< 0.000) 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 (0.61) is considered to be significantly less at the 0.05 level of significance.

P_d^{disc}: BLIND GRID versus LITTORAL. Using the example data above to compare probabilities of detection in the discrimination stage, 80 out of 100 and 8 out of 33 emplaced ordnance items were correctly discriminated as such in open water testing. Those four values are used to calculate a test statistic of 32.01. Since the test statistic is greater than the critical value of 3.84, the smaller discrimination stage detection rate (0.24) is considered to be significantly less at the 0.05 level of significance.

P_d^{res}: OPEN WATER versus LITTORAL. 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 0.10 level of significance.

P_d^{disc}: OPEN WATER versus LITTORAL. 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 two discrimination stage detection rates are considered to be significantly different at the 0.10 level of significance. While a significant result does not prove a cause and effect relationship exists between the change in survey area and change in performance, it does indicate that the ability of Demonstrator X to correctly discriminate seems to have been degraded by features of the littoral area relative to results from the open water using the same system.

APPENDIX B. USAEC BASED SYSTEM DETECTION SUMMARIES

TABLE B-1. AMEC/3DGEO SYSTEM DETECTION SUMMARY

Metric	Overall	By Projectile Caliber				
		40-mm	60-mm	81-mm	105-mm	155-mm
Blind grid						
Response stage						
P_d	65.5%	51.7%	31.0%	55.2%	89.7%	100.0%
P_d lower 90% confidence	60.0%	38.4%	19.7%	41.7%	78.4%	92.4%
P_{fp}	34.5%					
P_{fp} lower 90% confidence	29.7%					
P_{ba}	7.4%					
Discrimination stage						
P_d	60.0%	51.7%	17.2%	55.2%	86.2%	89.7%
P_d lower 90% confidence	54.4%	38.4%	8.6%	41.7%	74.3%	78.4%
P_{fp}	21.8%					
P_{fp} lower 90% confidence	17.8%					
P_{ba}	1.8%					
Littoral region						
Response stage						
P_d	29.7%	34.5%	6.9%	31.0%	24.1%	51.7%
P_d lower 90% confidence	24.7%	22.6%	1.8%	19.7%	14.0%	38.4%
P_{fp}	14.4%					
P_{fp} lower 90% confidence	11.0%					
BAR m^{-2}	0.019					
Discrimination stage						
P_d	29.0%	34.5%	6.9%	27.6%	24.1%	51.7%
P_d lower 90% confidence	24.0%	22.6%	1.8%	16.8%	14.0%	38.4%
P_{fp}	13.2%					
P_{fp} lower 90% confidence	10.0%					
BAR m^{-2}	0.016					
Response stage noise level: 0.09						
Recommended discrimination threshold: 3.6						

TABLE B-2. AMEC/3DGEO EFFICIENCY AND REJECTION RATES

	Efficiency	False-Positive Rejection Rate	Background Alarm Rejection Rate
Blind Grid			
At operating point	0.92	0.37	0.75
With no loss of P_d	1.00	0.13	0.50
Littoral Region			
At operating point	0.98	0.08	0.14
With no loss of P_d	1.00	0.04	0.11

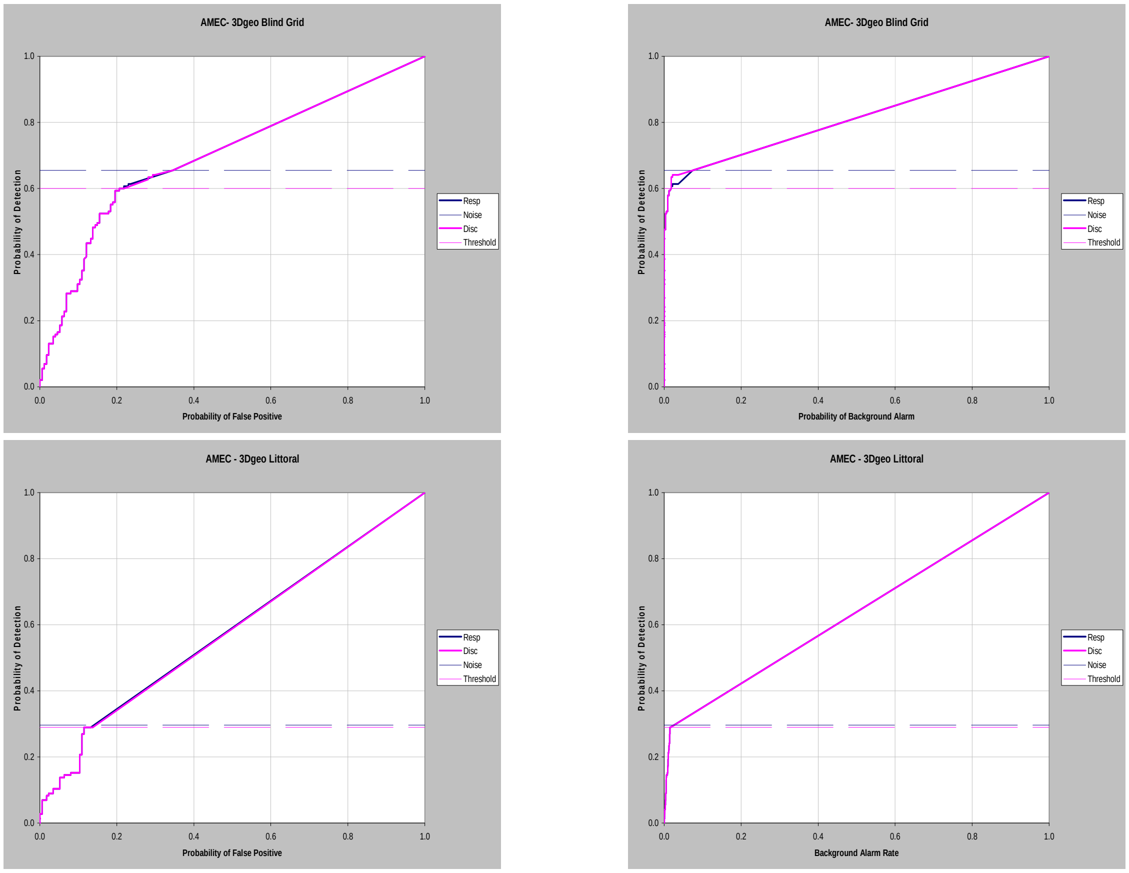


Figure B-1. AMEC/3Dgeo ROC curves.

TABLE B-3. CTC/FOERSTER SYSTEM DETECTION SUMMARY

Metric	Overall	By Projectile Caliber					
		40-mm	**60-mm**	**81-mm**	**105-mm**	**155-mm**	**8-inch**
Open Water							
Response Stage							
P_d	33.1%	34.5%	3.4%	17.2	62.1%	45.7%	33.3%
P_d Lower 90% Conf	28.2%	22.6%	0.4%	8.6%	48.5%	34.0%	9.3%
P_{fp}	12.3%						
P_{fp} Lower 90% Conf	9.4%						
BAR m^2	0.009						
Discrimination Stage							
P_d	33.1%	34.5%	3.4%	17.2%	62.1%	45.7%	33.3%
P_d lower 90% confidence	28.2%	22.6%	0.4%	8.6%	48.5%	34.0%	9.3%
P_{fp}	11.3%						
P_{fp} lower 90% confidence	8.5%						
BAR m^{-2}	0.009						

TABLE B-3. (CONT'D)

Metric	Overall	By Projectile Caliber					
		40-mm	60-mm	81-mm	105-mm	155-mm	8-inch
Littoral Region							
Response Stage							
P_d	15.9%	24.1%	0.0%	6.9%	3.4%	44.8%	
P_d Lower 90% Conf	12.0%	14.0%	0.0%	1.8%	0.4%	31.9%	
P_{fp}	6.9%						
P_{fp} Lower 90% Conf	4.5%						
BAR m^2	0.019						
Discrimination Stage							
P_d	14.5%	24.1%	0.0%	6.9%	3.4%	37.9%	
P_d Lower 90% Conf	10.8%	14.0%	0.0%	1.8%	0.4%	25.7%	
P_{fp}	6.9%						
P_{fp} Lower 90% Conf	4.5%						
Deeper Water							
Response Stage							
P_d	55.2%					55.2%	
P_d Lower 90% Conf	41.7%					41.7%	
Discrimination Stage							
P_d	55.2%					55.2%	
P_d Lower 90% Conf	41.7%					41.7%	
Response Noise Level	4						
Disc. Threshold	4						
Blind Grid							
Response Stage							
`P_d	56.6%	65.5%	6.9%	27.6%	82.8%	100.0%	
P_d Lower 90% Conf	50.9%	51.9%	1.8%	16.8%	70.3%	92.4%	
P_{fp}	28.2%						
P_{fp} Lower 90% Conf	23.7%						
P_{ba}	4.0%						
Discrimination Stage							
P_d	55.2%	65.5%	6.9%	27.6%	75.9%	100.0%	
P_d Lower 90% Conf	49.5%	51.9%	1.8%	16.8%	62.8%	92.4%	
P_{fp}	24.7%						
P_{fp} Lower 90% Conf	20.5%						
P_{ba}	4.0%						
Response Noise Level	0.5						
Disc. Threshold	1.5						

TABLE B-4. CTC/FOERSTER EFFICIENCY AND REJECTION RATES

	Efficiency	False-Positive Rejection Rate	Background Alarm Rejection Rate
Blind Grid			
At operating point	0.98	0.12	0.00
With no loss of P_d	1.00	0.12	0.00
Open Water			
At operating point	1.00	0.08	0.10
With no loss of P_d	1.00	1.00	1.00
Littoral Region			
At operating point	0.91	0.00	0.04
With no loss of P_d	1.00	0.00	0.04

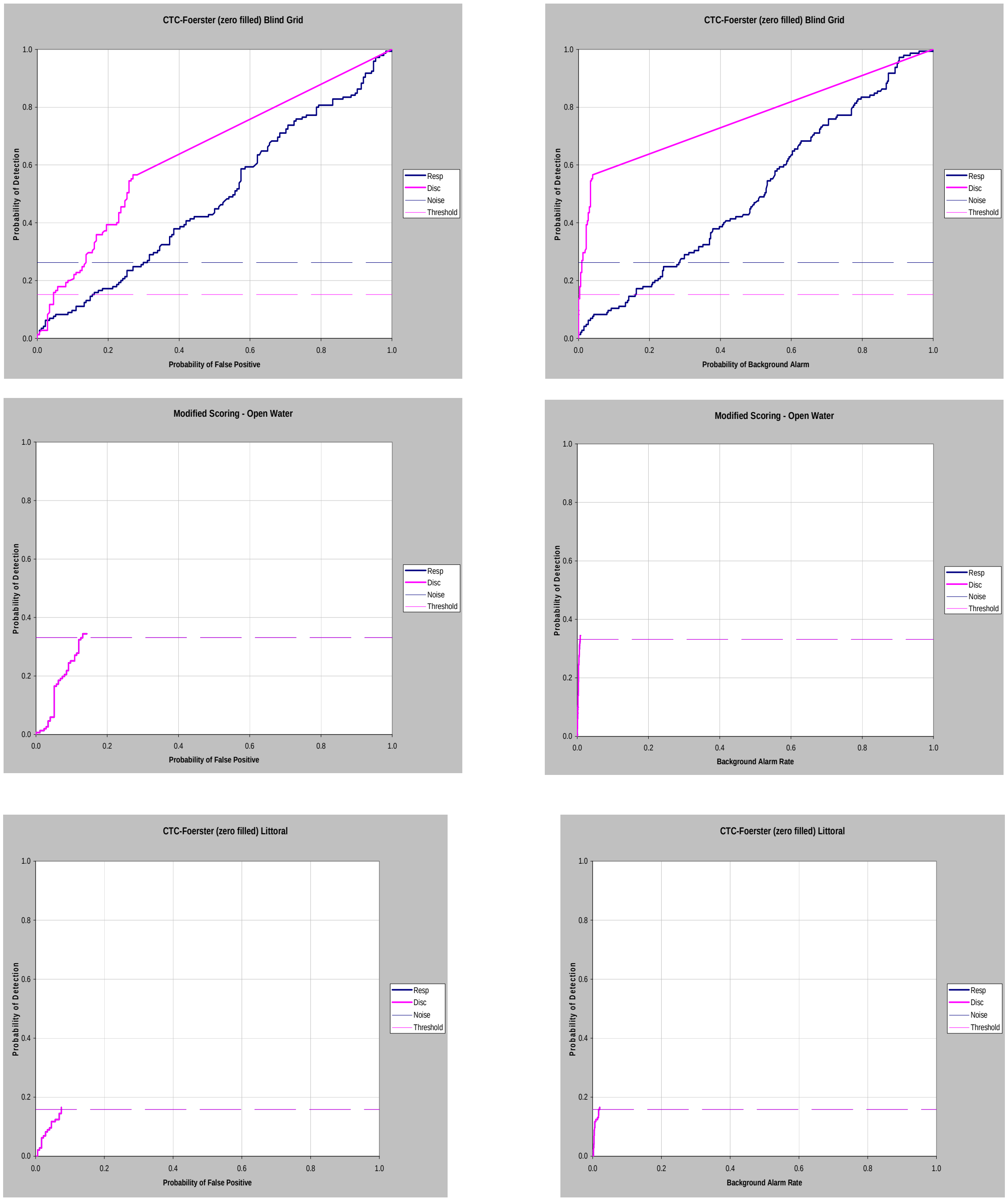


Figure B-2. CTC/Foerster ROC curves.

TABLE B-5. GEOPHEX SYSTEM DETECTION SUMMARY

Metric	Overall	By Projectile Caliber					
		40-mm	60-mm	81-mm	105-mm	155-mm	8-inch
Blind Grid							
Response Stage							
P_d	93.8%	79.3%	93.1%	96.6%	100.0%	100.0%	
P_d Lower 90% Confidence	90.4%	66.5%	82.7%	87.2%	92.4%	92.4%	
P_{fp}	73.6%						
P_{fp} Lower 90% Confidence	68.8%						
P_{ba}	7.7%						
Discrimination Stage							
P_d	82.8%	58.6%	79.3%	93.1%	89.7%	93.1%	
P_d Lower 90% Confidence	78.1%	45.0%	66.5%	82.7%	78.4%	82.7%	
P_{fp}	25.9%						
P_{fp} Lower 90% Confidence	21.5%						
P_{ba}	2.8%						
Open Water							
Response Stage							
P_d	68.2%	69.0%	65.5%	75.9%	79.3%	60.0%	33.3%
P_d Lower 90% Confidence	62.9%	55.5%	51.9%	62.8%	66.5%	47.8%	9.3%
P_{fp}	57.6%						
P_{fp} Lower 90% Confidence	52.9%						
BAR m^{-2}	0.051						
Discrimination Stage							
P_d	57.3%	69.0%	62.1%	62.1%	58.6%	45.7%	16.7%
P_d Lower 90% Confidence	51.9%	55.5%	48.5%	48.5%	45.0%	34.0%	1.7%
P_{fp}	18.2%						
P_{fp} Lower 90% Confidence	14.8%						
BAR m^{-2}	0.027						
Littoral Region							
Response Stage							
P_d	37.9%	27.6%	34.5%	41.4%	34.5%	51.7%	
P_d Lower 90% Confidence	32.6%	16.8%	22.6%	28.8%	22.6%	38.4%	
P_{fp}	32.2%						
P_{fp} Lower 90% Confidence	27.5%						
BAR m^{-2}	0.061						
Discrimination Stage							
P_d	30.3%	27.6%	27.6%	31.0%	24.1%	41.4%	
P_d Lower 90% Confidence	25.3%	16.8%	16.8%	19.7%	14.0%	28.8%	
P_{fp}	9.8%						
P_{fp} Lower 90% Confidence	7.0%						
BAR m^{-2}	0.034						
Deeper Water							
Response Stage							
P_d	75.9%					75.9%	
P_d Lower 90% Confidence	62.8%					62.8%	
Discrimination Stage							
P_d	75.9%					75.9%	
P_d Lower 90% Confidence	62.8%					62.8%	
Response Stage Noise Level: 20							
Recommended Discrimination Threshold: 5							

TABLE B-6. GEOPHEX EFFICIENCY AND REJECTION RATES

	Efficiency	False-Positive Rejection Rate	Background Alarm Rejection Rate
Blind Grid			
At operating point	0.88	0.65	0.64
With no loss of P_d	1.00	0.21	0.00
Open Water			
At operating point	0.84	0.68	0.47
With no loss of P_d	1.00	0.01	0.01
Littoral			
At operating point	0.80	0.70	0.44
With no loss of P_d	1.00	0.38	0.18

TABLE B-7. GEOPHEX CHI-SQUARE SIGNIFICANCE TEST RESULTS

Metric	Overall	By Projectile Caliber				
		40 mm	60 mm	81 mm	105 mm	155 mm
Blind Grid - Open Water Comparison						
P_d^{res}	Sig	Not	Sig	Sig	Sig	Sig
P_d^{disc}	Sig	[b]	Not	Sig	Sig	Sig
P_{fp}^{res}	Sig					
P_{fp}^{disc}	Sig					
Efficiency	Not					
R_{fp}	Not					
Sig = Significant Not = Not Significant						
[b]No test – Discrimination in the Open Water area is better than in the Blind Grid.						

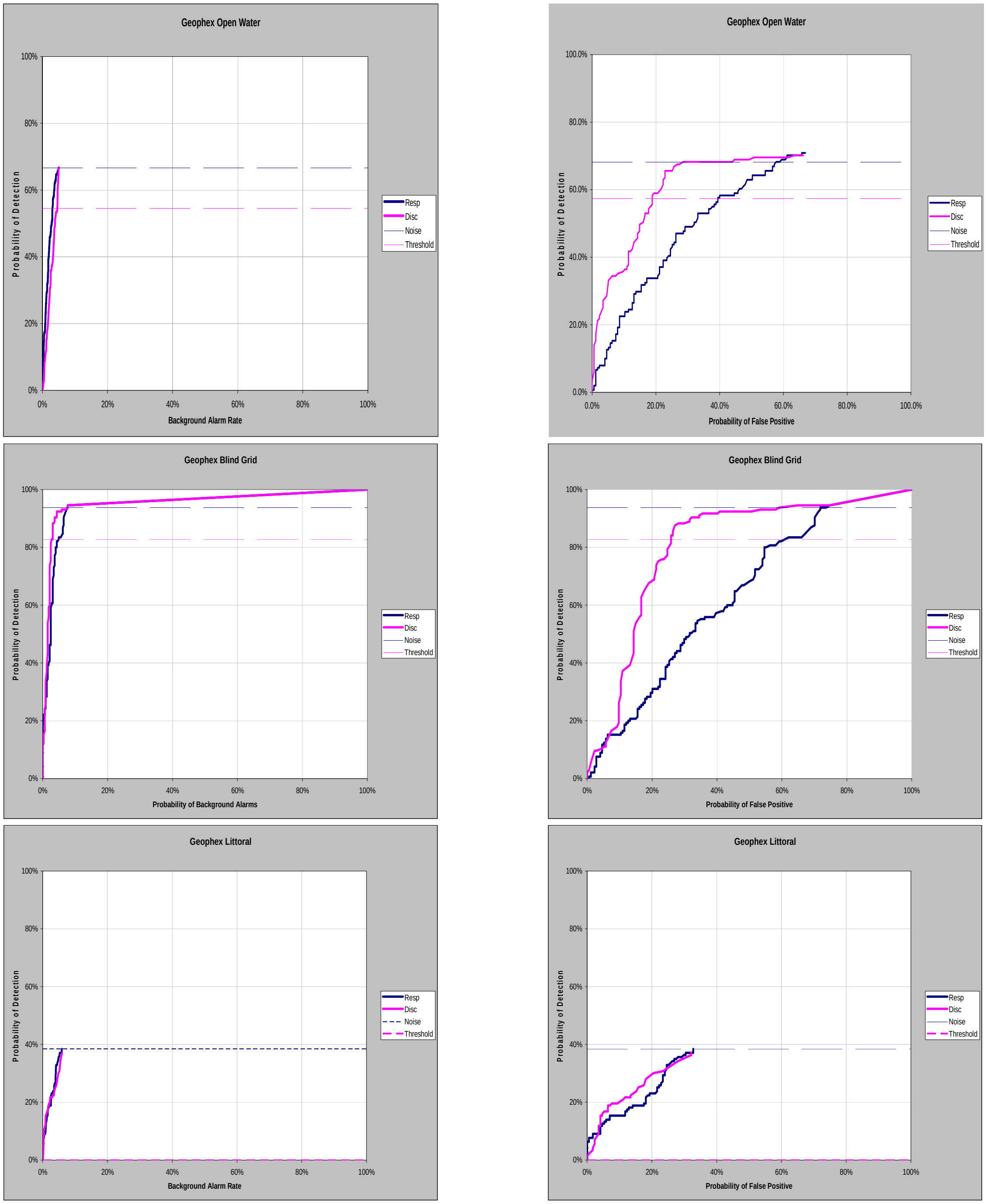


Figure B-3. Geophex ROC curves.

TABLE B-8. IT JEWELL MM SYSTEM DETECTION SUMMARY

MM Metric	Overall	By Projectile Caliber					
		40-mm	60-mm	81-mm	105-mm	155-mm	8-inch
Blind grid							
Response stage							
P_d	51.0%	51.7%	48.3%	48.3%	44.8%	62.1%	
P_d lower 90% confidence	45.4%	38.4%	35.1%	35.1%	31.9%	48.5%	
P_{fp}	56.3%						
P_{fp} lower 90% confidence	51.2%						
P_{ba}	46.5%						
Discrimination stage							
P_d	34.5%	34.5%	24.1%	31.0%	27.6%	55.2%	
P_d lower 90% confidence	29.3%	22.6%	14.0%	19.7%	16.8%	41.7%	
P_{fp}	37.4%						
P_{fp} lower 90% confidence	32.5%						
P_{ba}	29.2%						
Open water							
Response stage							
P_d	1.9%	3.4%	3.4%	0.0%	3.4%	0.0%	0.0%
P_d lower 90% confidence	0.7%	0.4%	0.4%	0.0%	0.4%	0.0%	0.0%
P_{fp}	2.5%						
P_{fp} lower 90% confidence	1.2%						
BAR m^{-2}	0.022						
Discrimination stage							
P_d	1.9%	3.4%	3.4%	0.0%	3.4%	0.0%	0.0%
P_d lower 90% confidence	0.7%	0.4%	0.4%	0.0%	0.4%	0.0%	0.0%
P_{fp}	2.5%						
P_{fp} lower 90% confidence	1.2%						
BAR m^{-2}	0.021						
Littoral region							
Response stage							
P_d	4.1%	6.9%	0.0%	3.4%	0.0%	10.3%	
P_d lower 90% confidence	2.2%	1.8%	0.0%	0.4%	0.0%	3.9%	
P_{fp}	2.9%						
P_{fp} lower 90% confidence	1.4%						
BAR m^{-2}	0.029						
Discrimination stage							
P_d	4.1%	6.9%	0.0%	3.4%	0.0%	10.3%	
P_d lower 90% confidence	2.2%	1.8%	0.0%	0.4%	0.0%	3.9%	
P_{fp}	2.9%						
P_{fp} lower 90% confidence	1.4%						
BAR m^{-2}	0.029						
Deeper water							
Response stage							
P_d	0.0%					0.0%	
P_d lower 90% confidence	0.0%					0.0%	
Discrimination stage							
P_d	0.0%					0.0%	
P_d lower 90% confidence	0.0%					0.0%	
Response stage noise level: 52620							
Recommended discrimination threshold: 5.099							

TABLE B-9.   IT JEWELL 882 SYSTEM DETECTION SUMMARY

G882 Metric	Overall	By Projectile Caliber					
		40-mm	60-mm	81-mm	105-mm	155-mm	8-inch
Blind grid							
Response stage							
P_d	29.0%	24.1%	24.1%	31.0%	31.0%	34.5%	
P_d lower 90% confidence	24.0%	14.0%	14.0%	19.7%	19.7%	22.6%	
P_{fp}	25.3%						
P_{fp} lower 90% confidence	21.0%						
P_{ba}	19.7%						
Discrimination stage							
P_d	19.3%	20.7%	13.8%	6.9%	24.1%	31.0%	
P_d lower 90% confidence	15.1%	11.2%	6.2%	1.8%	14.0%	19.7%	
P_{fp}	20.1%						
P_{fp} lower 90% confidence	16.2%						
P_{ba}	15.1%						
Open water							
Response stage							
P_d	0.6%	0.0%	0.0%	3.4%	0.0%	0.0%	0.0%
P_d lower 90% confidence	0.1%	0.0%	0.0%	0.4%	0.0%	0.0%	0.0%
P_{fp}	2.5%						
P_{fp} lower 90% confidence	1.2%						
BAR m^{-2}	0.022						
Discrimination stage							
P_d	0.6%	0.0%	0.0%	3.4%	0.0%	0.0%	0.0%
P_d lower 90% confidence	0.1%	0.0%	0.0%	0.4%	0.0%	0.0%	0.0%
P_{fp}	2.5%						
P_{fp} lower 90% confidence	1.2%						
BAR m^{-2}	0.021						
Littoral region							
Response stage							
P_d	3.4%	13.8%	0.0%	3.4%	0.0%	0.0%	
P_d lower 90% confidence	1.7%	6.2%	0.0%	0.4%	0.0%	0.0%	
P_{fp}	5.7%						
P_{fp} lower 90% confidence	3.6%						
BAR m^{-2}	0.038						
Discrimination stage							
P_d	3.4%	13.8%	0.0%	3.4%	0.0%	0.0%	
P_d lower 90% confidence	1.7%	6.2%	0.0%	0.4%	0.0%	0.0%	
P_{fp}	5.7%						
P_{fp} lower 90% confidence	3.6%						
BAR m^{-2}	0.038						
Deeper water							
Response stage							
P_d	0.0%					0.0%	
P_d lower 90% confidence	0.0%					0.0%	
Discrimination stage							
P_d	0.0%					0.0%	
P_d lower 90% confidence	0.0%					0.0%	
Response stage noise level: 52625							
Recommended discrimination threshold: 4.18							

TABLE B-10. IT JEWELL EFFICIENCY AND REJECTION RATES

	Efficiency		**False-Positive Rejection Rate**		**Background Alarm Rejection Rate**	
Blind grid						
	G882	**MM**	**G882**	**MM**	**G882**	**MM**
At operating point	0.67	0.68	0.20	0.34	0.23	0.37
With no loss of P_d	1.00	1.00	0.00	0.00	0.03	0.00
Open water						
	G882	**MM**	**G882**	**MM**	**G882**	**MM**
At operating point	1.00	1.00	0.00	0.00	0.06	0.04
With no loss of P_d	1.00	1.00	0.40	0.20	0.65	0.56
Littoral region						
	G882	**MM**	**G882**	**MM**	**G882**	**MM**
At operating point	1.00	1.00	0.00	0.00	0.00	0.00
With no loss of P_d	1.00	1.00	0.10	0.80	0.37	0.55

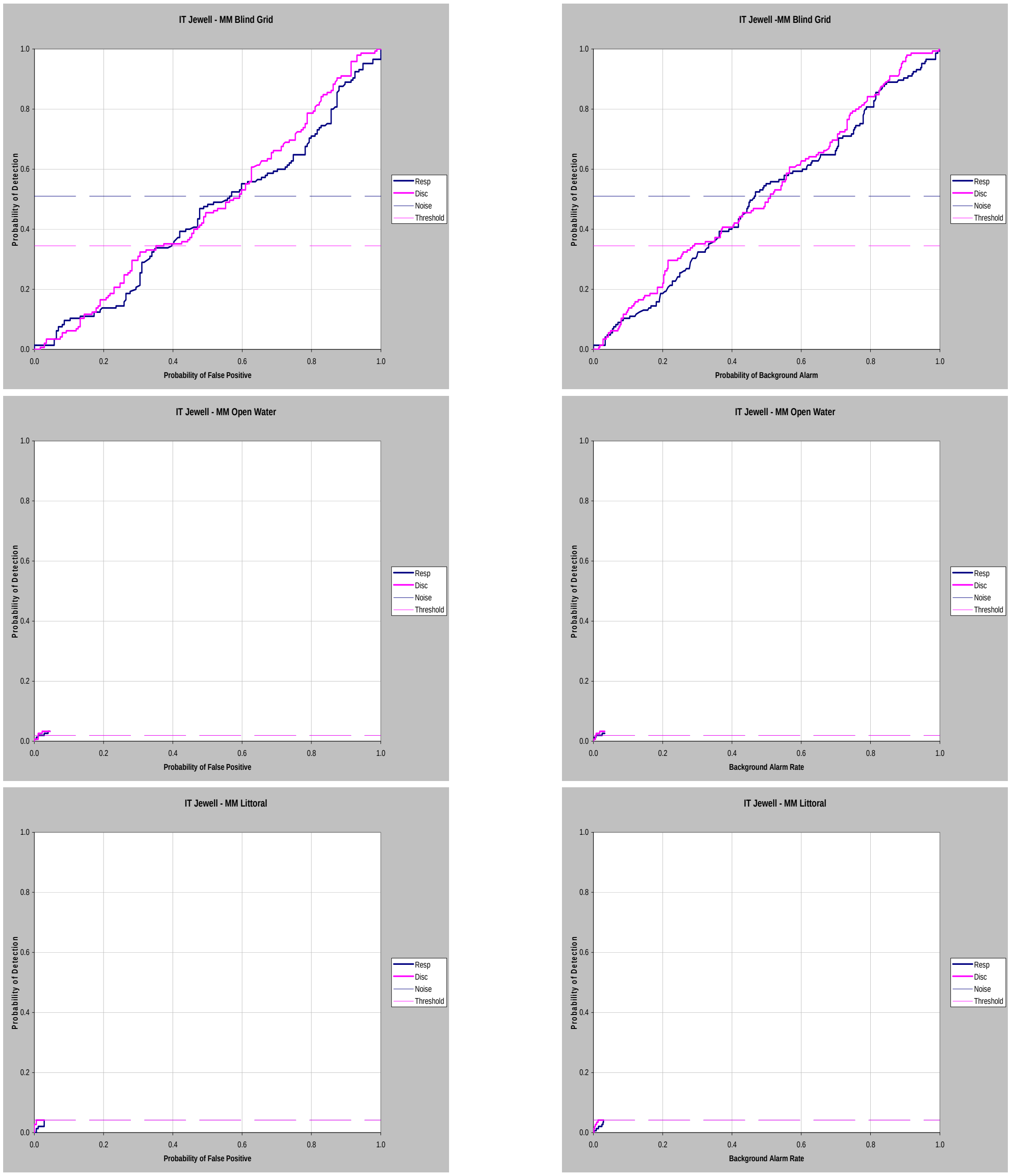


Figure B-4. IT Jewell MM system ROC curves.

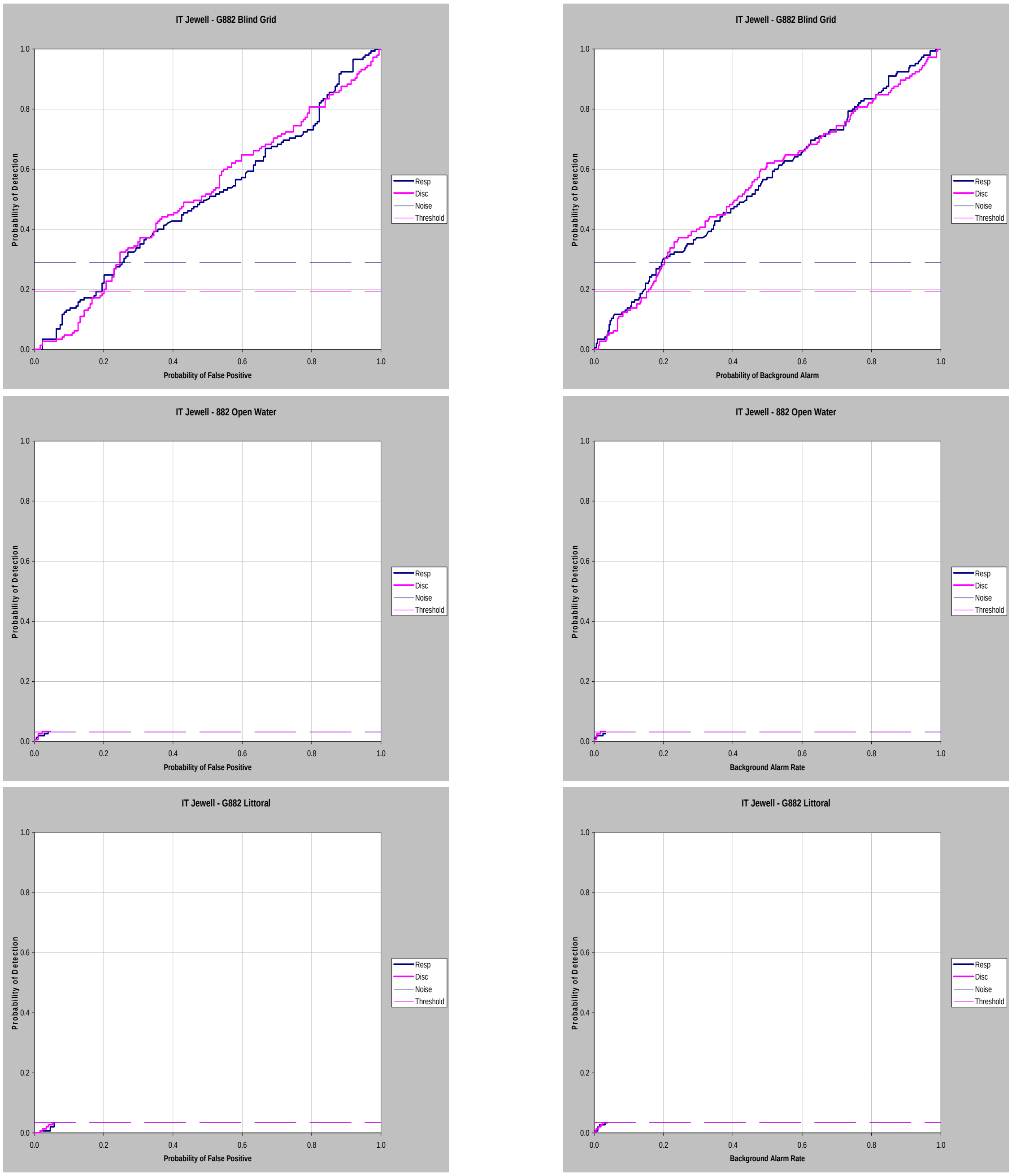


Figure B-5. IT Jewell 882 system ROC curves.

TABLE B-11. NAEVA/XTECH SYSTEM DETECTION SUMMARY

Metric	Overall	By Projectile Caliber					
		40-mm	60-mm	81-mm	105-mm	155-mm	8-inch
Blind grid							
Response stage							
P_d	95.2%	96.6%	93.1%	93.1%	100.0%	93.1%	
P_d lower 90% confidence	92.0%	87.2%	82.7%	82.7%	92.4%	82.7%	
P_{fp}	92.0%						
P_{fp} lower 90% confidence	88.6%						
P_{ba}	5.8						
Discrimination stage							
P_d	45.5%	27.6%	44.8%	37.9%	55.2%	62.1%	
P_d lower 90% confidence	39.9%	16.8%	31.9%	25.7%	41.7%	48.5%	
P_{fp}	56.9%						
P_{fp} lower 90% confidence	51.8%						
P_{ba}	1.8						
Open water							
Response stage							
P_d	54.8%	62.1%	51.7%	41.4%	62.1%	57.1%	50.0%
P_d lower 90% confidence	49.3%	48.5%	38.4%	28.8%	48.5%	44.9%	20.1%
P_{fp}	54.2%						
P_{fp} lower 90% confidence	49.4%						
BAR m^{-2}	0.016						
Discrimination stage							
P_d	31.8%	24.1%	24.1%	20.7%	55.2%	37.1%	16.7%
P_d lower 90% confidence	27.0%	14.0%	14.0%	11.2%	41.7%	26.1%	1.7%
P_{fp}	30.0%						
P_{fp} lower 90% confidence	25.8%						
BAR m^{-2}	0.007						
Littoral region							
Response stage							
P_d	Test area not surveyed						
P_d lower 90% confidence							
P_{fp}							
P_{fp} lower 90% confidence							
BAR m^{-2}							
Discrimination stage							
P_d	Test area not surveyed						
P_d lower 90% confidence							
P_{fp}							
P_{fp} lower 90% confidence							
BAR m^{-2}							
Deeper water							
Response stage							
P_d	24.1%					24.1%	
P_d lower 90% confidence	14.0%					14.0%	
Discrimination stage							
P_d	24.1%					24.1%	
P_d lower 90% confidence	14.0%					14.0%	
Response stage noise level: 0.55							
Recommended discrimination threshold: 1.5							

TABLE B-12. NAEVA/XTECH EFFICIENCY AND REJECTION RATES

	Efficiency	False-Positive Rejection Rate	Background Alarm Rejection Rate
Blind Grid			
At operating point	0.48	0.38	0.68
With no loss of P_d	1.00	0.38	0.68
Open Water			
At operating point	0.58	0.45	0.58
With no loss of P_d	1.00	0.45	0.58
Littoral Region			
At operating point	Test area not surveyed		
With no loss of P_d			

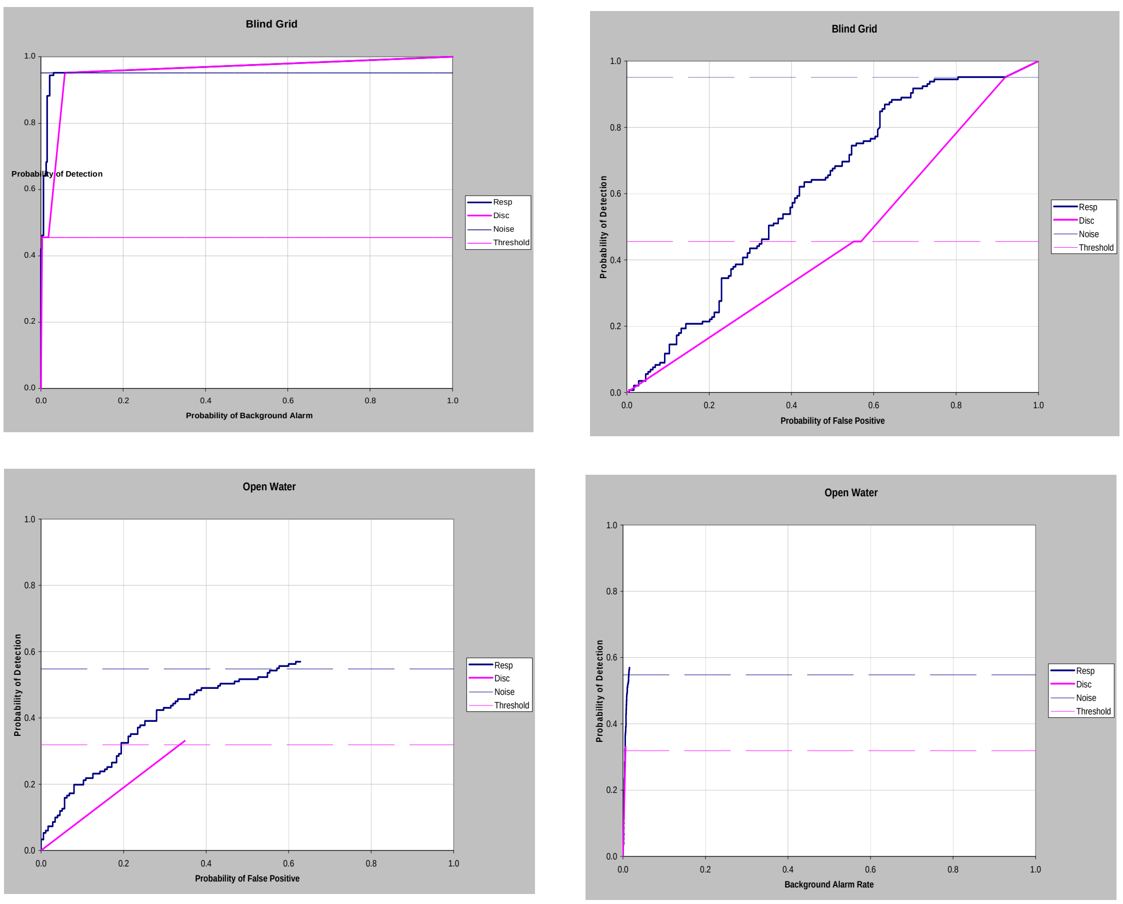


Figure B-6. NAEVA\XTECH ROC curves.

TABLE B-13. NAEVA/3DGEO SYSTEM DETECTION SUMMARY

Metric	Overall	By Projectile Caliber					
		40-mm	60-mm	81-mm	105-mm	155-mm	8-inch
Blind grid							
Response stage							
P_d	95.2%	96.6%	93.1%	93.1%	100.0%	93.1%	
P_d lower 90% confidence	92.0%	87.2%	82.7%	82.7%	92.4%	82.7%	
P_{fp}	92.0%						
P_{fp} lower 90% confidence	88.6%						
P_{ba}	5.8						
Discrimination stage							
P_d	45.5%	27.6%	44.8%	37.9%	55.2%	62.1%	
P_d lower 90% confidence	39.9%	16.8%	31.9%	25.7%	41.7%	48.5%	
P_{fp}	56.9%						
P_{fp} lower 90% confidence	51.8%						
P_{ba}	1.8						
Littoral region							
Response stage							
P_d	17.9%	13.8%	6.9%	27.6%	17.2%	24.1%	
P_d lower 90% confidence	13.9%	6.2%	1.8%	16.8%	8.6%	14.0%	
P_{fp}	15.5%						
P_{fp} lower 90% confidence	12.0%						
BAR m^{-2}	0.012						
Discrimination stage							
P_d	12.4%	13.8%	0.0%	27.6%	6.9%	13.8%	
P_d lower 90% confidence	9.0%	6.2%	0.0%	16.8%	1.8%	6.2%	
P_{fp}	5.7%						
P_{fp} lower 90% confidence	3.6%						
BAR m^{-2}	0.004						
Response stage noise level: 160							
Recommended discrimination threshold: 1							

TABLE B-14. NAEVA/3DGEO SYSTEM EFFICIENCY AND REJECTION RATES

	Efficiency	**False-Positive Rejection Rate**	**Background Alarm Rejection Rate**
Blind Grid			
At operating point	0.52	0.64	1.00
With no loss of P_d	1.00	0.64	1.00
Littoral Region			
At operating point	0.69	0.63	0.63
With no loss of P_d	1.00	0.63	0.63

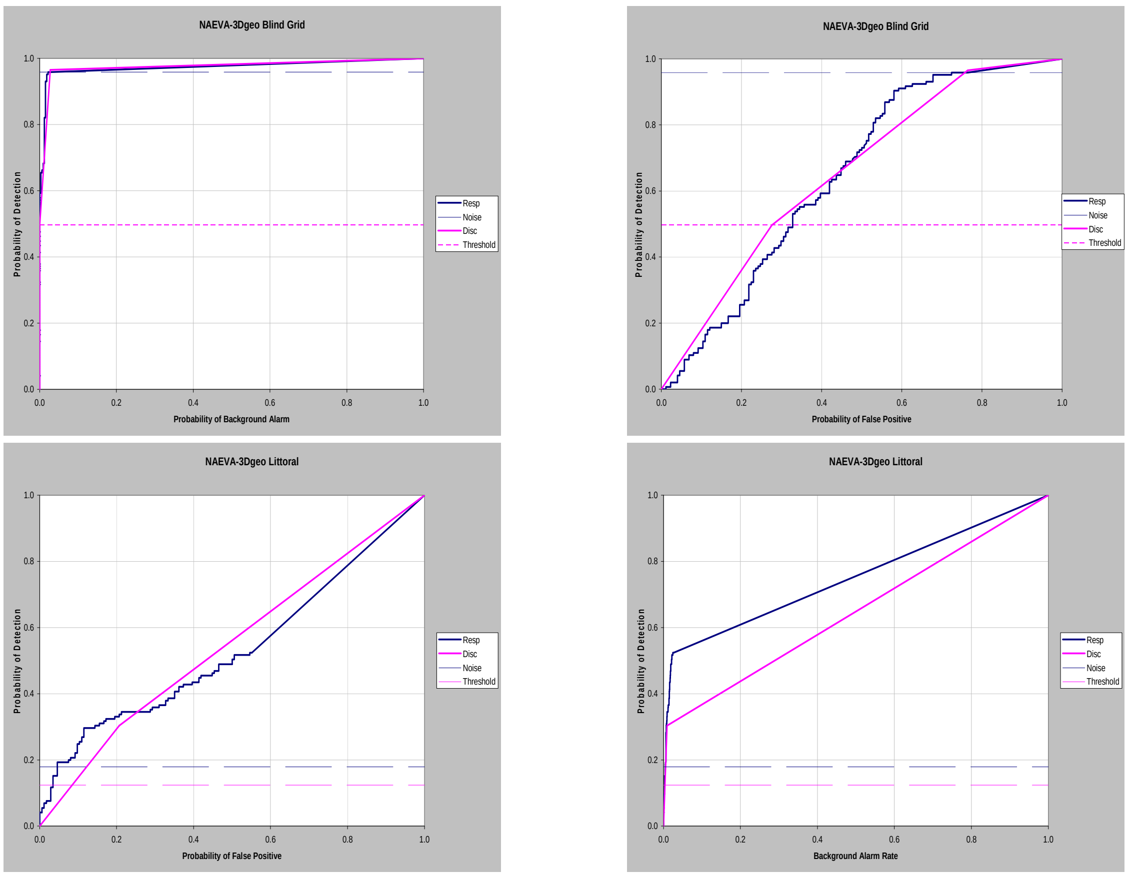


Figure B-7. NAEVA / 3DGEO ROC curves.

TABLE B-15. TETRA TECH SYSTEM DETECTION SUMMARY

Metric	Overall	By Projectile Caliber					
		40-mm	60-mm	81-mm	105-mm	155-mm	8-inch
Blind Grid							
Response Stage							
P_d	47.6%	58.6%	41.4%	51.7%	37.9%	48.3%	
P_d Lower 90% Confidence	42.0%	45.0%	28.8%	38.4%	25.7%	35.1%	
P_{fp}	37.9%						
P_{fp} Lower 90% Confidence	33.1%						
P_{ba}	44.6%						
Discrimination Stage							
P_d	42.1%	44.8%	37.9%	44.8%	37.9%	44.8%	
P_d Lower 90% Confidence	36.6%	31.9%	25.7%	31.9%	25.7%	31.9%	
P_{fp}	32.8%						
P_{fp} Lower 90% Confidence	28.1%						
P_{ba}	39.7%						
Open Water							
Response Stage							
P_d	5.7%	10.3%	6.9%	3.4%	6.9%	2.9%	0.0%
P_d Lower 90% Confidence	3.5%	3.9%	1.8%	0.4%	1.8%	0.3%	0.0%
P_{fp}	7.4%						
P_{fp} Lower 90% Confidence	5.1%						
BAR m^{-2}	0.020						
Discrimination Stage							
P_d	3.8%	3.4%	6.9%	3.4%	3.4%	2.9%	0.0%
P_d Lower 90% Confidence	2.0%	0.4%	1.8%	0.4%	0.4%	0.3%	0.0%
P_{fp}	4.9%						
P_{fp} Lower 90% Confidence	3.1%						
BAR m^{-2}	0.012						
Littoral Region							
Response Stage							
P_d	23.4%	10.3%	3.4%	17.2%	44.8%	41.4%	
P_d Lower 90% Confidence	18.9%	3.9%	0.4%	8.6%	31.9%	28.8%	
P_{fp}	8.0%						
P_{fp} Lower 90% Confidence	5.5%						
BAR m^{-2}	0.030						
Discrimination Stage							
P_d	19.3%	10.3%	3.4%	17.2%	34.5%	31.0%	
P_d Lower 90% Confidence	15.1%	3.9%	0.4%	8.6%	22.6%	19.7%	
P_{fp}	6.9%						
P_{fp} Lower 90% Confidence	4.5%						
BAR m^{-2}	0.028						
Deeper Water							
Response Stage							
P_d	6.9%					6.9%	
P_d Lower 90% Confidence	1.8%					1.8%	
Discrimination Stage							
P_d	6.9%					6.9%	
P_d Lower 90% Confidence	1.8%					1.8%	
Response Stage Noise Level: 0.1 for all areas							
Recommended Discrimination Threshold: Blind Grid =3 Open Water, Littoral and Deeper Water =1							

TABLE B-16. TETRA TECH EFFICIENCY AND REJECTION RATES

	Efficiency	**False-Positive Rejection Rate**	**Background Alarm Rejection Rate**
Blind grid			
At operating point	0.88	0.14	0.11
With no loss of P_d	1.00	0.02	0.03
Open water			
At operating point	0.67	0.33	0.40
With no loss of P_d	1.00	0.13	0.02
Littoral region			
At operating point	0.82	0.14	0.06
With no loss of P_d	1.00	0.00	0.01

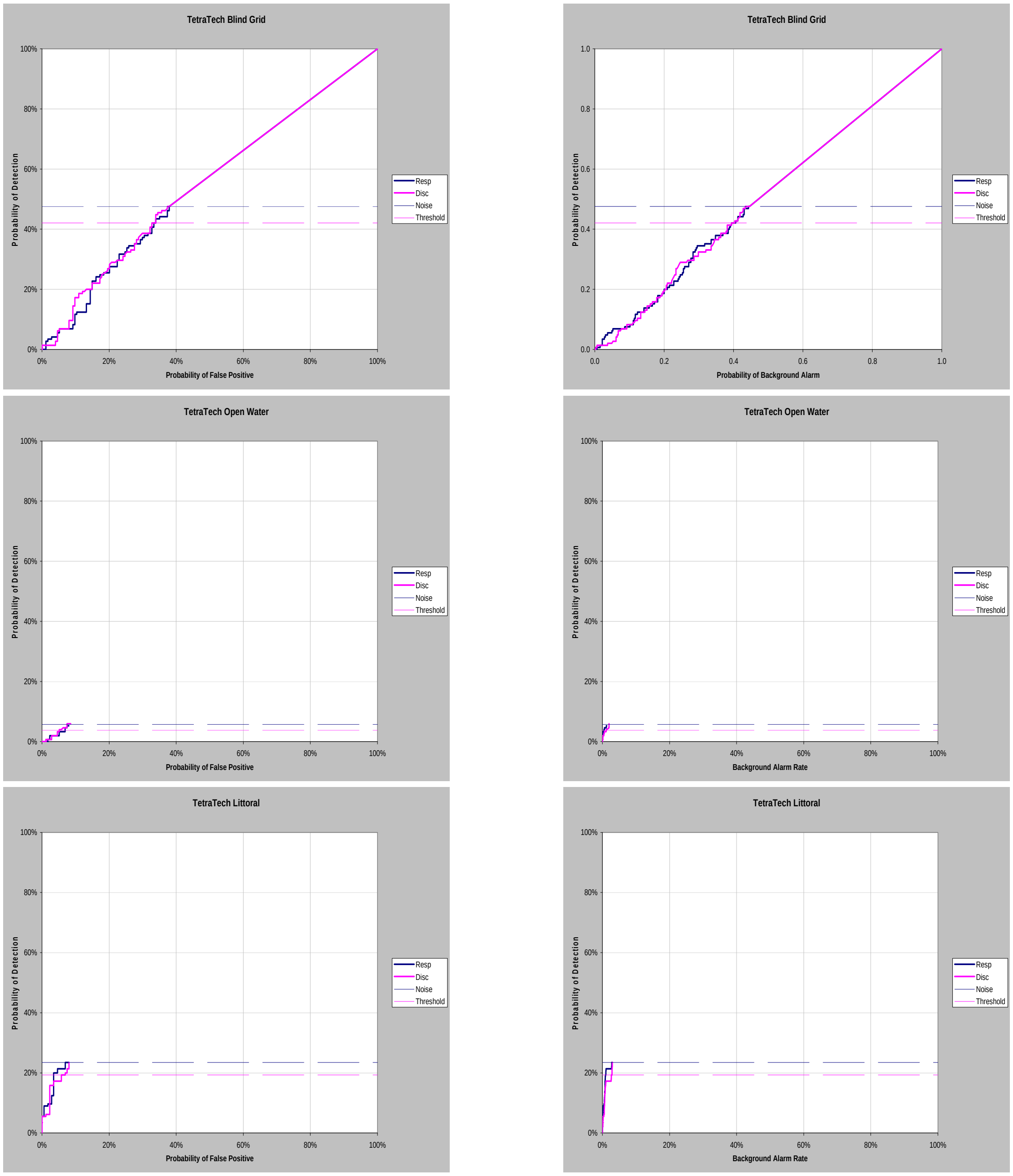


Figure B-8. Tetra Tech ROC curves.

APPENDIX C.   REFERENCES

1. Broad Agency Announcement W91ZLK-04-R-0001, Unexploded Ordnance (UXO) Detection and Discrimination Demonstrations, 18 November 2004.

2. Validation of Detection Systems (VDS) Test Program Report, Mare Island Naval Shipyard (MINS), Vallejo, California. Prepared by the Environmental Chemical Corporation under Navy Contract No. N62742-98-D-1809, CTO 0001, 7 July 2000.

3. Email, SWDS Scoring, 6 July 2006, sent from Mr. Timothy Deigan (timothy.deigan@tteci.com) to Mr. Gary Rowe (gary.rowe@atc.army.mil).

4. Andrews, A., V. George, T. Altshuler, and M. Mulqueen, "Results of the Countermine Task Force Mine Detection Demonstration at Fort A.P. Hill, Virginia, March 18-22, 1996," IDA paper P-3192, July 1996.

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

APPENDIX D. ABBREVIATIONS

ABS	=	acrylonitrile butadiene styrene
APG	=	Aberdeen Proving Ground
ATC	=	U.S. Army Aberdeen Test Center
BAA	=	Broad Agency Announcement
CD-R	=	compact disc - recordable
CTC	=	Concurrent Technologies Corporation
DGPS	=	Differential Global Positioning System
EM	=	electromagnetic
EMI	=	electromagnetic induction
EQT	=	Environmental Quality Technology
ERDC	=	U.S. Army Corps of Engineers, Engineer Research and Development Center
ESTCP	=	Environmental Security Technology Certification Program
FFT	=	Fast Fourier Transform
GPS	=	Global Positioning System
GT	=	ground truth
HFIS	=	high frequency imaging sonar
HP	=	high power
IDA	=	Institute for Defense Analyses
MEC	=	munitions and explosives of concern
METDC	=	Military Environmental Technology Demonstration Center
MFSBP	=	multifrequency subbottom profile
MGS	=	magnetic gradiometer system
MINS	=	Mare Island Naval Shipyard
MTADS	=	Multi-Sensor Towed Array Detection System
MX	=	mixed metals
NRL	=	Naval Research Laboratory
POC	=	point of contact
PVC	=	polyvinyl chloride
QA	=	quality assurance
ROC	=	receiver operating characteristics
Rx	=	receiver
SERDP	=	Strategic Environmental Research and Development Program
USAEC	=	U.S. Army Environmental Command
UTM	=	Universal Transverse Mercator
UXO	=	unexploded ordnance
VDS	=	Validation of Detection Systems

APPENDIX E.   DISTRIBUTION LIST

DTC Project No. 8-CO-160-UXO-016

Addressee	No. of Copies
Commander U.S. Army Aberdeen Test Center ATTN: CSTE-DTC-AT-SL-E (Mr. Gary Rowe)	1
(Library)	1
CSTE-DTC-AT-CS-M Aberdeen Proving Ground, MD 21005-5055	1
Commander U.S. Army Environmental Command ATTN: IMAE-ATT (Ms. Kimberly Watts) Aberdeen Proving Ground, MD 21010-5401	1
Booz Allen Hamilton ATTN: Mr. Patrick McDonnell 4692 Millennium Drive, Suite 200 Belcamp, MD 21017-1535	1
SERDP/ESTCP ATTN: Ms. Anne Andrews 901 North Stuart Street, Suite 303 Arlington, VA 22203	1
Defense Technical Information Center 8725 John J. Kingman Road, Suite 0944 Fort Belvoir, VA 22060-6218	PDF

Secondary hard-copy distribution is controlled by Commander, U.S. Army Environmental Command, ATTN: IMAE-ATT.