USER GUIDE AND DESIGN DOCUMENT

Supporting Operation of the Technology Developed and Demonstrated in ESTCP Project MR-201714: AUV-Based Structural Acoustic Sonars for Underwater Buried UXO Detection and Classification

B.H. Houston, D. Raudales, H. Simpson, A. Berdoz,
Z.J. Waters, and T. Yoder,
Naval Research Laboratory, Washington, DC 20375

J. A. Bucaro,
Excet, Inc., Springfield, VA 22510

A. Moskalenko and A. Sarkissian,
Sotera KeyW, Columbia, MD 21046

May 27, 2021

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
27/05/2021	ESTCP User's Guide	

4. TITLE AND SUBTITLE

Demonstration of AUV-Based Structural Acoustic Look-Down and Side-Look Sonars for Underwater Buried UXO Detection and Classification: User's Guide and Design Document

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Brian Houston, D. Raudales, H. Simpson, A. Berdoz, Z.J. Waters, and T. Yoder
Naval Research Laboratory

Joseph Bucaro
EXCET, Inc.

A. Moskalenko and A. Sarkissian,
Sotera KeyW

5d. PROJECT NUMBER
MR-201714

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Naval Research Laboratory
4555 Overlook Ave
Code7130
Washington, DC 20375

8. PERFORMING ORGANIZATION REPORT NUMBER

MR-201714

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Environmental Security Technology Certification Program 4800 Mark Center Drive, Suite 16F16
Alexandria, VA 22350-3605

10. SPONSOR/MONITOR'S ACRONYM(S)

ESTCP

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

MR-201714

12. DISTRIBUTION/AVAILABILITY STATEMENT

DISTRIBUTION STATEMENT A. Approved for public release: distribution unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

This document outlines the major steps to be carried out by a test organization tasked with prosecuting the target detection and classification activities of an UXO underwater clean-up operation using the technology demonstrated in ESTCP Project MR-201714: AUV-Based Structural Acoustic Sonars for Underwater Buried UXO Detection and Classification.

15. SUBJECT TERMS

Autonomous Underwater Vehicle, AUV-Based Structural Acoustic Look-Down Sonar, Side-Look Sonar, Underwater Buried UXO Detection, Unexploded ordnance

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Brian Houston
UNCLASS	UNCLASS	UNCLASS	UNCLASS	25	**19b. TELEPHONE NUMBER** *(Include area code)* 202-404-3840

TABLE OF CONTENTS

LIST OF ACRONYMS

AUV	Autonomous Underwater Vehicle
DAS	Digital Acquisition System
DVL	Doppler Velocity Log
FIR	Finite Input Response
FOG	Fiber Optic Gyro
GPS	Global Positioning System
INS	Inertial Navigation System
ITAR	International Traffic in Arms Regulations
IVER	Integration Vehicle Engineering Release
LH	Lincoln Hat
LRD	Launch and Recovery Device
MATLAB	Matrix Laboratory
MOB	Mobilized
NRL	Naval Research Laboratory
RHIB	Rigid Hull Inflatable Boat
ROC	Receiver Operating Characteristics
RVM	Relevance Vector Machine
SAS	Synthetic Aperture Sonar
SNR	Signal-to-Noise Ratio
TVR	Transmitter Voltage Response
UXO	Unexploded Ordnance

LIST OF FIGURES

USER GUIDE AND DESIGN DOCUMENT

This document outlines the major steps to be carried out by a test organization tasked with prosecuting the target detection and classification activities of an UXO underwater clean-up operation using the technology demonstrated in ESTCP Project MR-201714: *AUV-Based Structural Acoustic Sonars for Underwater Buried UXO Detection and Classification.*

We assume that the detection and classification process is to be carried for both proud and buried UXO articles using the Reliant side-look system for the former and the down-look system for the latter. The activities to be carried out are organized into the following categories: (A) Site Preparation; (B) AUV - Based Structural Acoustic Sonar Preparation; (C) Seeded Site Training Data Collection; (D) Classification Algorithm(s) Training; (E) Test Data Collection; (F) Generation of Detection Display Maps; (G) Generation of Classification Call Lists; and (H) Generation of Final "Dig List."

(A) Site Preparation

Prior to establishing the target training field, an evaluation should be made as to whether adequate information is available regarding the sediment and bottom properties at the cleanup site. In particular, it would be useful to have bathymetric, backscatter, side-scan and sub-bottom sonar data. The bathymetric data provides measurements of water depth that, when processed, are used to map the seafloor topography. Backscatter and side-scan sonar data provide images that support characterization of surface sediment texture and roughness. Sub-bottom sonar data can be used to characterize site geology including the composition of near subsurface sediments as well as any interfering outcropping structure. CR Environmental, Inc. performed such a sediment classification survey for Bluefin Robotics Corporation and the U.S. Naval Research Laboratory in support of the latter's ESTCP Demonstration Program prior to target deployments and autonomous underwater vehicle (AUV) exercises at the NRL Box 1A Demonstration Test Site in Nahant Bay, MA. An acoustic survey consisting of multi-beam bathymetry and sub-bottom profiling was performed on 26 April, 2018, and an underwater video survey and sediment grab sampling were conducted on 27 April 2018 to ground truth the acoustic data. These exercises and their results represent good examples of important supporting information and how it can be obtained.

A major part of site preparation involves the design and establishment of the proud and buried target training fields. It is assumed that the training area will be established in a region included in the remediation site to support the aim of matching the training and testing environments. If this is not possible, a training site should be found having comparable sediment characteristics. After deployment of training targets, the training field environment should be thoroughly surveyed with a side-scan imaging sonar to locate clutter and determine ground truth locations for targets. In the case of buried training targets, a team of divers should be employed to manually determine target burial depths and orientations and to place acoustic reflectors for locating targets using the high frequency side-scan sonar. Scattering monuments (e.g. Lincoln Hats) are introduced to facilitate target localization and mitigate navigational uncertainties stemming from the lack of GPS underwater. It is important that surveying methods used to localize targets (i.e. side-scan, divers, etc.) are completed close to the date of the training exercise measurements, as environmental factors can cause targets to naturally migrate (both laterally and in depth) over time.

The targets used for training should include a reasonable number of each UXO type expected to be found at the test site as well as a number of non-UXO targets (e.g. concrete blocks, paint cans, etc.) if they are thought likely to be present there as well. The number of training targets to be used has not been rigorously established. However, we know that the more variations existing in the sediment properties and target characteristics, the greater statistical sampling is required to achieve good classification and "dig list" performance. The same targets could be used for establishing the proud and buried training fields assuming that the first is established and characterized completely before the second is established and characterized. The buried training field should be roughly fifty meters by fifty meters in extent with the targets buried nominally ten meters apart over a range of pitch angles from zero to thirty degrees. All targets could be at the same azimuthal angles and buried at a range of depths from flush to that of the deepest UXO thought to be present at the clean-up site. Four-inch diameter, thirty-nine inch high stainless steel pipes (with a twelve inch spike welded on one end) which serve as ground truth buried target markers should be stuck in the sediment positioned over each of the buried targets. Although probably not needed, the pipes can also be used in the proud field as well alongside each UXO target. Pipe images generated by the high-resolution imaging sonar are subsequently used to localize the training field targets. The proud target training field should be about one hundred and fifty meters by one hundred meters with the proud targets placed on average about twenty meters apart. Final preparation for the proud and buried fields includes deployment of the Lincoln hat navigation monuments. These are five-hundred-pound steel objects designed to have strong, omni-directional, clearly recognizable frequency-angle returns from any of the sonars deployed in the exercises (the down-look, the Reliant side-look, and the high-resolution Solstice imaging sonar). The monuments have a twelve inch tall eight inch diameter solid steel cylinder sitting atop a two-and-one-half inch thick twenty inch diameter steel disc (ergo the name Lincoln hat). These monuments should be placed proud in both training fields in a regular pattern about twenty meters apart in the buried field and fifty meters apart in the proud field. All target localization information will be referenced to these monument sites. All of the pipe reflectors are removed after completing the high-resolution target localization scans with a high frequency imaging sonar such as the one used in our program *i.e.* the Sonardyne International Limited Solstice sonar mounted on a 12inch Bluefin AUV. Use of the Lincoln hat monuments has been critical to the success of the demonstration program and will be so in future efforts. At the present time, they are important in two respects. (1) They allow the accurate transformation of single-line, limited aspect echo data to 360 degree multi-aspect acoustic color maps. And (2), accurate target localization is carried out *with respect to the Lincoln hat monument locations.* If this were an actual remediation clean-up operation, divers would find the specified targets by first physically locating the Lincoln hats and then moving from them to the targets following the localization information generated from the tests.

We believe that the a-priori geolocation accuracy of the Lincoln Hats (LH) themselves can be improved. If one could achieve an improvement allowing LH geolocation accuracy approaching one meter, this would in turn provide absolute target localization to this accuracy. Note that the use of a bottom fiducial (e.g. Lincoln Hat) for doing underwater target geolocation is a commonly used technique particularly as it applies to very deep surveys. Unlike deep ocean problems, the shallow water surveys in this work also support an underwater navigation process where the vehicle never loses DVL "bottom lock". Maintaining bottom lock minimizes INS drift and allows the INS to maintain the highest navigation accuracy. Leveraging recent work at NRL, further refinement involves an approach where several dives are used to collect multiple geolocations to improve the overall LH location accuracy. Further, the improved process involves the use of

modern “re-navigation” algorithms that support the careful assimilation of multiple GPS location “hits” together with “back propagation” of the vehicle navigation information as the AUV vehicle obtains these multiple GPS hits through multiple dives and multiple passes on the LH markers. It should be noted that recent efforts show that it may be possible to eliminate the use of the LH markers altogether by relying simply on the topographical structure of the bottom to carry out the geolocation. This approach leverages techniques that are similar to what is used in terrain-based navigation for sonars.

(B) AUV - Based Structural Acoustic Sonar Preparation

To operate the down look sonar and Reliant, a support vessel needs to be identified. The support vessel should have a large enough operations deck to house the launch and recovery device (LRD) for the AUV. Our current LRD requires an A-frame on the stern of the support vessel that is used to raise and lower a ramp into the water. The support vessel should have a freeboard less than 7’. Typically, the support vessels are at least 50’ long. Larger support vessels (larger than 100’) would require a RHIB (rigid hull inflatable boat) or small boat to be used during the recovery process. The smaller support vessels can usually maneuver to recover the AUV without the need for a small boat or RHIB.

The support vessel must also be relatively quiet with engines off and only the generator running. If the on-board generator is too noisy, a deck-top generator can be installed to minimize acoustic noise in the water during AUV sonar operation.

The AUVs are relatively compact; however, some support materials are required. The AUV system can be housed, fully assembled for shipping, in the cart section of the LRD, or separated into five components and placed in individual shipping containers. In general, the AUVs (down-look sonar and Reliant) use the same three sections of the AUV, i.e. the drive section (tail cone), energy section, and data acquisition system, (DAS). The receiver and source sections are different for Reliant and down-look sonar. In addition to the AUV components, operation of the AUVs requires support equipment and spare parts. The spare parts include a spare drive motor or a spare tail cone section, plus a litany of spare cables and other components. These spare components and topside support equipment typically fit into two or three Ropak containers. In addition to the cart and ramp sections of the LRD, five assembly carts usually ship with the AUV along with a spreader bar to crane the AUV. Crane lifting of the AUV is only done dockside and should never be attempted at-sea.

Once all the equipment is on location, the support vessel is mobilized (MOB) by attaching the LRD. Various adapter plates have been made to fit a variety of support vessels; however, each new support vessel will require modifications in order to mount the LRD securely. In addition to the LRD, the top side equipment (mostly electronics and antennas) are mounted on the support vessel. It typically takes several hours to mobilize a support vessel. A deck-top generator is usually installed to minimize the noise from the support ship during acoustic operations. The radiated noise is usually assessed during a pre-measurement visit to the support vessel. To MOB usually requires a dockside crane or a ship board crane.

Once the support vessel is mobilized, the AUV can be assembled and prepared for operations. The AUV preparations procedures are as follows:

1) Assemble AUV in Reliant or down-look sonar configuration
2) Perform deck-top system integration tests
3) Perform dockside in-water integration tests including initial ballasting and vehicle weight trimming
4) Carry out at-sea dynamic AUV control tuning and final weight ballasting and weight trimming
5) Carry out at-sea final AUV system integration test

The mobilization and initial configuration integrated system testing usually takes one to two days. Afterward, the AUV system is ready for operations with either down-look sonar or Reliant sonars. Switching between down-look sonar and Reliant configurations can be easily done in one day, repeating the AUV preparation procedures.

(C) Seeded Site Training Data Collection

As discussed above, prior to the structural acoustic data collection exercises, a high resolution scanning sonar (e.g. the Solstice sonar) is used to localize each UXO training target by imaging the pipe reflectors. These positions are registered as referenced to the Lincoln hat monuments. This target localization information will be used when training the classification algorithms.

The actual acoustic training data (and later the test data itself) is collected using the two structural acoustic sonars. The scans are carried out over linear flight paths in east-west, north-south, and two diagonal directions ninety degrees apart. The Reliant side-looking sonar scans the proud target field and the down-looking sonar scans the buried target field. The down-look sonar flights are separated by two meters and the Reliant flights by twenty meters. At both ends of each flight path, approximately fifty meters of distance is allowed for executing AUV turn arounds.

Timing and time gates

The down-look sonar electronics and software onboard the down-look sonar vehicle are designed to excite the projector every 40 ms with a 4 ms chirped broadband electrical waveform. The side-look sonar electronics and software onboard the Reliant vehicle are designed to excite the projector every 300 ms also with a 4 ms chirped broadband electrical waveform. Prior to carrying out SAS processing, the measured signal at each receiver will be time windowed to remove the bottom bounce from directly below the vehicle (and everything arriving prior to it) and the water surface reflection (with everything arriving after it). Further, the incident pressure spectrum will be removed from the received signals by dividing the latter by the product of the spectrum of the electrical exciting waveform and the transmitting voltage response (TVR) projector curve (i.e. matched filter processed) as determined in our calibration measurements and reported in our Final Report. If this former source calibration is no longer appropriate, new source calibration measurements should be collected as described in our report. This processing component is a three

step process: 1) Fourier transform the raw time data into the frequency domain, 2) apply a bandpass filter to remove data outside of the experimental frequency band, and 3) divide the received pressure by the source strength (this is a measurement of the source response).

Navigation and positioning

Vehicle navigation/positioning information is obtained through a navigation solution based on measurements from GPS, an on-board Doppler Velocity Log (DVL), and a fiber optic gyro (FOG)-based inertial navigation system (IXBLUE) which together provide vehicle position, yaw, pitch, and roll time histories with sufficient accuracies so as not to limit spatial sonar resolution. The requirement is $\lambda/8$ precision which would require position registration on the order of 0.01m (0.4 inches). This information is used to establish the actual positions versus time of the 64 receivers and projector required for robust SAS processing of the acoustic data.

Data acquisition

The side-look and down-look DAS are virtually identical. The data acquisition system accepts digitized data at a 62.5 kHz rate from the receiver elements (digital hydrophones) and has a dynamic range of 16 bits. The data is sent to a main controller housed in the interior of a watertight section of the AUV where each hydrophone signal is organized into a larger data file that is broadcast through the system to be either analyzed or recorded. The digital hydrophones have a low pass digital filter (132 points finite input response or FIR) starting at 26 kHz with 80dB rejection at 31.25 kHz. The high pass filtering is accomplished by a set of three-pole elliptical filters, the first set at 250Hz and the second set at 1.5 kHz. The DAS accepts external triggers and synchronous triggers and can be slaved to the distance travelled.

(D) Classification Algorithm(s) Training

Assuming the training site environmental acoustics, sediment characteristics, and clutter characteristics are representative of what would be experienced at the clean-up site, all that would be required is that the training data contain a sufficient structural acoustic sampling of the UXO types and their burial circumstances and that these are typical of those expected to be encountered at the clean-up site. The goal in developing the classifier is to design a feature space which would effectively separate the UXO targets from both the known and unknown clutter using information extracted from the scattering signature. Due to the typically limited availability of UXOs in the training exercise (which in our demonstration program included only 12 UXOs), a relevance vector machine or RVM - rather than a neural network which traditionally requires larger training data sets - was selected to be more suitable for classification. Additionally, the dimensionality of the feature vector was limited to two in order to reduce the risk of overtraining which could result in instabilities due to the "curse of dimensionality" related to the imbalance between the measurement and feature space dimensionalities. We considered four features and three different pairs of feature parameters for the trained RVM algorithm. The features are: (A) time correlation; (B) beam highlight width; (C) acoustic color symmetry correlation; and (D) detection map correlation. The three feature set pairs for which we trained the RVM algorithm are: A-B; D-A; and C-A. As we discuss below, we have the most confidence in the feature set C-A. In the following, we refer to feature A as the correlation feature and feature C as the symmetry feature. Our strategy in curating the feature architecture was to examine the existing available data and determine what factors differentiated the UXOs from the clutter in the temporal and spectral domain of the target strengths

(acoustic color). For example, in observing the free-field response of the UXOs both in the NRL large pool facility and from numerical models, their response in the NRL sediment pool, the data collected in our previous SERDP funded effort in the Gulf, and the shakedown and training data collected in the Boston Harbor, we concluded that the beam aspect return (beam aspect defined as the incident angle at which the AUV sonar when running parallel to the cylindrical axis of the target is closest to the target) would yield the highest broadband scattering levels from the UXOs and also provide the best opportunity for classification. The parabolic-like shape of the time return (see Figure 1) is attributed to the flat-panel-like perspective of the pseudo-cylindrical targets and can be used to differentiate UXO from many manmade targets and more irregularly shaped clutter objects. A correlation feature was designed which applied a 2D cross-correlation to the temporal target strength (acoustic color) using templates derived from the beam aspect return of the at-sea training data, which included all three classes of UXO targets prosecuted in the NRL Demonstration Program.

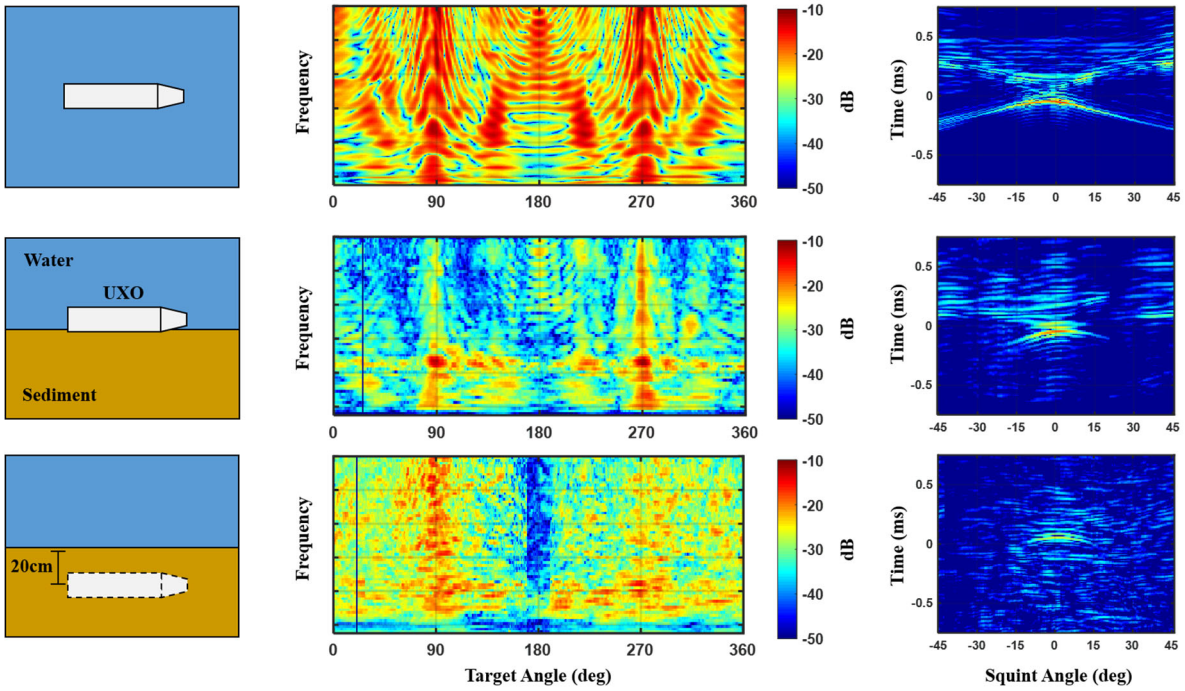

Figure 1. Comparison of target strengths for Howitzer class UXO target in free-field (top row, measured at NRL structural acoustics pool facility), proud state (middle row, measured at 28m range by side-look in blind test) and buried state (bottom row, measured at 3m range by down-look in blind test). Multi-aspect spectral response is shown in middle column and temporal response for beam aspect return (designated at 90° and 270° of the target angle) is shown in right column. The at-sea measurements for the proud and buried UXO are processed using a 0.4m synthetic aperture resolution.

Operating in the temporal domain rather than the frequency domain allows for the correlation feature to more uniquely capture both the shape and size of the target; it however requires high precision in target focusing. Further, the geometric orientation of the target must match that at which it is measured in the training template. The two angles of interest are the azimuthal angle in

the horizontal plane and the polar or pitch angle (for buried or partially buried targets). Having multiple vehicle headings in the measurement plan takes care of the azimuthal angle and ensures that a beam aspect return would be guaranteed for each deployed target. Regarding the polar or pitch angle, we make the assumption that buried UXO targets are found more often closer to a horizontal polar burial condition than a more vertical orientation. Thus, we include only a small range of pitch angles beyond horizontal burial. This assumption is reasonable to the extent that the area is not prone to significant weather events. If this assumption cannot be made, one needs to include the full 90 degree range of buried pitch angles in the UXO training field in steps of about 10 degrees.

The second feature used in classification consists of a novel symmetry feature which exploits the multi-aspect approach in processing and operates in concert with the correlation feature to further isolate the UXOs in the feature space. This feature, which was introduced in our research effort, is derived by performing a cross-correlation on two temporal scattering responses of the same target surveyed with an aspect separated by 180°, or on opposite sides of the target. The response from manmade targets especially UXOs generally contains broadband symmetry that provides unique differentiation from what is observed in scattering from much of the natural clutter such as rock formations which lack symmetry that may otherwise result in false alarms.

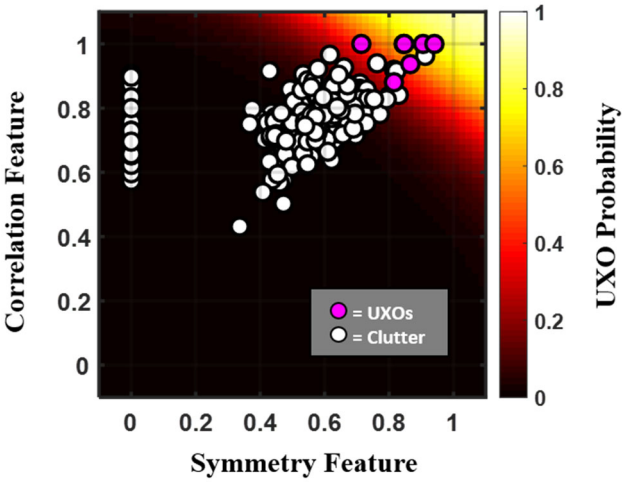

Figure 2. Two-dimensional feature space plot displaying symmetry and correlation feature values for UXOs and clutter deployed in our ESTCP Program buried training field. Distribution of UXO-like probability as computed through discriminative RVM training is presented in background heat map.

Once computed from the scattering data, training is accomplished by feeding the correlation and symmetry feature values into the Bayesian model of the RVM to generate a probabilistic distribution solution that optimizes separation between targets and clutter in the 2D feature space. As evident in the results presented in Fig. 1, detections are deemed UXO-like if there is both strong correlation and 180° symmetry in the target strengths (acoustic color). Training can be performed discriminatively using all classes of UXOs present in the training field and the clutter, both deployed and indigenous, detected in the buried training exercise. The UXO templates for the correlation feature are extracted from the at-sea training measurements. An advantage of incorporating the symmetry feature lies in its ability to remove clutter objects that contained specular glints which correlated well with the UXO templates but were only illuminated within a limited target aspect. These false alarms are displayed in the left-hand side of the feature space plot in Fig. 2 and are particularly challenging to eliminate through a single pass approach.

(E) Test Data Collection

The acoustic testing data is collected at the clean-up site using the two structural acoustic sonars sequentially. The scans are carried out over "mow-the-grass" linear flight paths in east-west, north-south, and two diagonal directions ninety degrees apart. The Reliant side-looking sonar scans the target field searching for UXO proud targets and the down-looking sonar scans the target field prosecuting potential buried (and proud) UXO. The down-look sonar flights are separated by two meters and the Reliant flights by twenty meters.

(F) Generation of Detection Display Maps

There are three principal data products generated by post-processing the raw data streams. These principal data products are: (1) Target Images; (2) Acoustic Color; and (3) Time-angle Plots. The first data product, the Target Image, is used with both down-look sonar and Reliant data collection exercises in generating the detection display maps. The latter two data products, i.e. (1) Acoustic Color and (2) Time-angle Plots, are used in the target classification process. In the case of down-look sonar, the target images are also used to generate supporting target information including expert classification calls, burial depth, and target orientation.

Conversion of Navigation and Acoustic Data

The vehicle output from each line measurement is converted into our universal format resulting in a simplified data structure that contains pertinent vehicle information including the position tracking as reported by the inertial navigation system (INS), altitude from the Doppler Velocity Log (DVL), the sound speed measurement and system metadata. The dynamic measurements for the yaw, pitch and roll of the vehicle are used in conjunction with the static vehicle geometry to register the source and receiver locations at each ping.

The sonar output from the projector (for down-look sonar, from each of the two transducers) is stored in binary format and de-convolved using the broadside free-field reference measured (in our case) in the NRL structural acoustics laboratory. For our program, the reference measurement was obtained by firing the exact waveform used at-sea with the vehicle placed in a large vibration isolated and temperature-controlled tank. The response was measured in the near-field using a high precision robotically controlled receiver array and projected into the far-field by applying acoustic holography. Measurements were collected for varying horizontal and vertical pitch angles of the vehicle, which are applied in the post-processing of target strengths to compensate for the aspect dependence of the source.

The raw output from the INS contains errors in reported vehicle dynamics that if unaddressed can propagate into the target strengths as artificial spectral "dropouts" that inhibit classifier performance. Our approach for improving the accuracy and smoothness of the vehicle navigation consisted of feeding the raw data from each line measurement into the DELPH INS software, a Kalman filter-based algorithm developed by Bluefin, resulting in a post-processed re-navigated solution. Additionally, noise in the vehicle altitude, as reported by the DVL, is reduced by realigning the measurement with the specular reflection off the seafloor as calculated from the sonar data. This empirical corrective process, made possible by the downward oriented source and

receiver array, results in higher fidelity SAS processing and ensured that beam-formed target responses were properly centered at the t = 0 bin.

Generation of Detection Surface

Having de-convolved the acoustic data and registered the source and receiver positions, the acoustic detection map for localizing targets is generated through synthetic aperture processing (SAS) as detailed in the following analysis. A Cartesian coordinate system is defined for each detection surface as having the x-axis pointing north, the y-axis pointing east and the positive z-axis pointing into the sediment. For the down-look system, a three-dimensional volumetric detection grid, which will later be compressed into a two-dimensional surface, is required to localize the buried targets. This grid is constructed using a uniform 15cm resolution in all three dimensions, out to a max range of 14m on each side of the vehicle projected along the entirety of the linear path. Vertical bounds for the volumetric grid are defined 15cm above the sediment and typically 75cm below the sediment. The latter bound can be increased if it is suspected that deeper UXO may be present at the clean-up site. For the side-look system, only one z coordinate is necessary i.e. that locating the sediment surface.

Once the grid is formed, a standard delay-and-sum beamforming algorithm is implemented to extract the real aperture target response:

$$p\left(\vec{r}_i, t_i, \theta_s\right) = \frac{1}{N_r} \sum_j \left|\vec{r}_{sn} - \vec{r}_i\right| \left|\vec{r}_{jn} - \vec{r}_i\right| p_{jn}\left(\left(|\vec{r}_{jn} - \vec{r}_i| + |\vec{r}_{sn} - \vec{r}_i|\right)/c + t_i\right) \quad . \tag{1}$$

Here, the real aperture response p is generated by steering the acoustic beam projected from the source at ping n_i, or squint angle θ_s, onto grid point r_i. This calculation is performed by applying a digital time delay of $\Delta t = \left(|\vec{r}_{jn} - \vec{r}_i| + |\vec{r}_{sn} - \vec{r}_i|\right)/c$ to each of the de-convolved receiver responses p_{jn} and coherently summing the response over all N_r receivers. The receiver and source positions at ping n are designated by r_{jn} and r_{sn}, respectively. Note that the response is multiplied by the two-way propagation distance to account for spherical spreading.

In order to generate the synthetic aperture response p_{sas}, the real aperture response is summed over a synthetic aperture consisting of N_p number of pings:

$$p_{SAS}\left(\vec{r}_i, t_i, \theta_s\right) = \frac{1}{N_p} \sum_n p\left(\vec{r}_i, t_i, \theta_s\right) \quad . \tag{2}$$

The synthetic aperture length, or number of pings, is range dependent and calculated using the following formula to achieve an azimuthal imaging resolution Δs, where Δd defines the average ping separation:

$$N_p = \frac{\lambda \left|\vec{r}_{sn} - \vec{r}_i\right|}{\Delta d \Delta s} \quad . \tag{3}$$

To improve runtime efficiency in generating the detection map, the aperture length is approximated using the center band wavelength λ_c. A SAS resolution of Δs = 40cm, comparable to the nominal UXO size, is chosen for the detector.

The center ping n_i used in forming the aperture dictates the angle θ_s at which the target is illuminated. Selecting the ping at the point of closest approach results in direct observation of the target, and hence selecting an off-angle ping is referred to as "squinting". The angular relation is determined from the following equation, in which r_{cpa} and n_{cpa} define the source location and ping number at the point of closest approach:

$$\theta_s = \arctan\left(\frac{\Delta d\left(n_i - n_{cpa}\right)}{\left|\left(\vec{r}_{cpa} - \vec{r}_i\right)\right|}\right) \quad . \tag{4}$$

The time delayed response at each grid point is computed for a squint range of ±45 degrees, dictated by the operating limits of the source directivity, with an angular spacing of Δ1 degree. This fine angular resolution is required to ensure that the narrow specular glint behavior of UXO targets is captured in the detection.

The real and synthetic aperture responses, p and p_{sas}, which are defined here in the context of detection, are later referred to as the temporal target strength and can be converted into the acoustic color, or spectral response, through their Fourier transform pairs, P and P_{sas}. The detector score σ is calculated by integrating the mean-square response across time window Δt_w, chosen to be 0.6ms, and angular window $\Delta\theta_w$, defined previously as 90 deg:

$$\sigma(\vec{r}_i) = \frac{1}{\Delta\theta_w \Delta t_w} \int_{-\Delta\theta_w/2}^{\Delta\theta_w/2} \int_{\Delta t_w/2}^{\Delta t_w/2} \left| p_{SAS}\left(\vec{r}_i, t, \theta\right)\right|^2 dt d\theta \, . \tag{5}$$

For down-look sonar, detection scores are generated at two separate frequency bands: 1) the low frequency MASA transducer band (3.5 - 8kHz) and 2) the high frequency ITC transducer band (10 - 24kHz). The lower frequency band penetrates further into the sediment and is included to facilitate buried object detection as well as to expand the overall bandwidth. The two sources are not collocated and thus the exact location for each source is used in calculating the near range time delays to ensure accuracy.

The computations involved in generating the detection scores are intensive, especially for down-look sonar due to the added vertical dimensionality required for buried object detection. In this case, one requires target strength computations for approximately 1,000,000 grid points per each measurement. The processing time is significantly reduced by parallelizing the computations on a high-performance computer using the MATLAB parallel processing toolbox.

Image Filtering for Detection Surface

For the down-look sonar, the 3D detection volume is compressed into a 2D detection surface by taking the maximum value of the detection score over the vertical axis, or depth, for each lateral

grid point and normalizing the result over the entire surface such that values can be reported in units of standard deviations. For Reliant, one simply takes the z = 0 value of the detection score.

The following describes the image filtering techniques applied to the detection surface in order to produce a meaningful result which can be interpreted for target localization. The first step applies a median filter to the image along the down-range component to compensate for the range dependence of both the environmental reverberation, which varies significantly for the down-look measurements, and the source beam pattern. The detection surface is then convolved with a disk filter, using a 25cm radius, to illuminate detections of sizes comparable to the targets. For the down-look sonar which utilizes two sources, the resulting low and high frequency detection surfaces are averaged together, weighted by the bandwidth of each source, to produce the filtered dual-band detection map. Fig. 3 provides an example comparison between the raw and filtered detection surface.

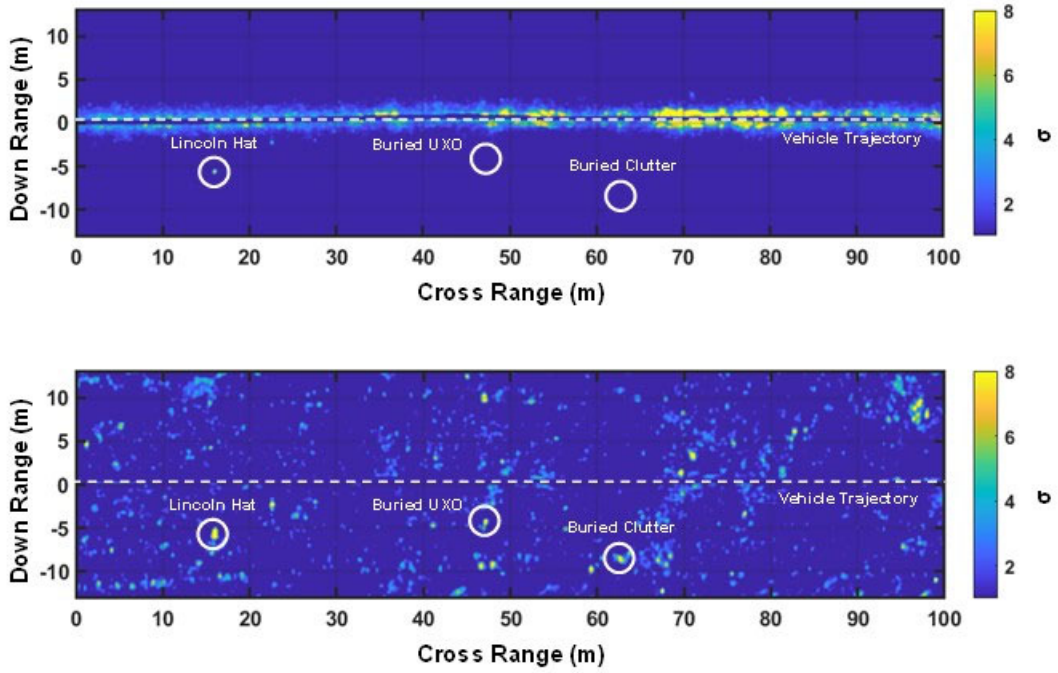

Figure 3. A comparison between the raw (above) and filtered (below) detection surfaces.

Prior to further analysis, the detection surface for each line is manually re-navigated to ensure high precision in multi-aspect target viewing. We implemented a correlation based approach in which a human expert translated, but did not rotate, each 2D detection surface such that clearly illuminated proud monuments and major indigenous features would become aligned on a universal grid. This process, demonstrated in Fig. 4, assumes the first line in the measurement is accurately navigated, a reasonable assumption given the monotonic behavior of vehicular drift, and performs alignment on the subsequent line which contains the most overlap in detection surfaces. The procedure is iterated until each line in the measurement has been re-navigated. Detection surface alignment in the testing or clean-up field is facilitated by the strategic placement of Lincoln Hats, which act as omnidirectional high intensity scattering monuments, such as to maximize the number of lines in which at least one monument is visible.

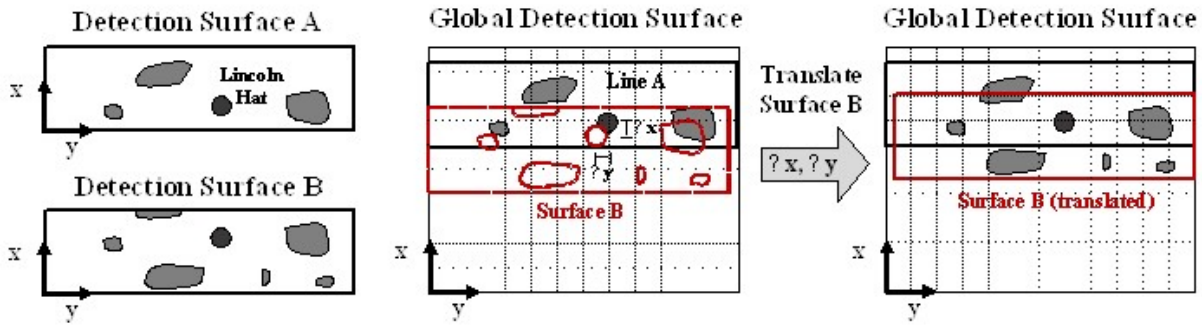

Figure 4. Cartoon illustrating re-navigation process.

Target Strength Extraction

For the down-look system, in which finer precision is required in aligning the target response than the detection grid resolution, the process of extracting target strengths is performed after generating the detection surface. The filtered dual-band surface is first converted into a binary image by applying a threshold of σ = 2.5, a value determined to guarantee detection of the low SNR targets according to the results from the training data. The binary image is then operated on using MATLAB's image processing toolbox to group clustered binary components together into defined objects which are stored for analysis. For each detected object above threshold, a localized high-resolution 3D volumetric grid (10cm resolution) is constructed and the detection score σ is computed for each grid point in the cube using the beamforming algorithm from eq. (1) with a reduced time window of 0.3ms. The best focus point is chosen by selecting the grid point which yields the maximum detector score and the target strength is extracted using the following relation:

$$TS(\omega_i, \theta_s) = 20 \log_{10} \left(\left| P_{SAS} \left(\vec{r}_i, \omega_i, \theta_s \right) \right| \right). \tag{8}$$

The synthetic aperture length used in calculating the beam-formed response here is now frequency dependent rather than approximated using the center-band frequency to produce a more exact result. The source directivity is accounted for in the target strength by calibrating the levels according to the horizontal and vertical source response taken from the earlier free-field measurement.

Implementing separate algorithms for the detector and target strength extraction is desired for two reasons: (1) to improve computational efficiency and (2) to decouple the imaging parameters. As previously discussed, generating detector scores using the down-look system although expensive does not, however, require the same fidelity as do the products that are used in classification, such as the target strength. This allows for simplifying assumptions so as long as they do not compromise the signal-to-noise ratio. Additionally, the imaging parameters, specifically the grid resolution and time window, used in extracting the target strengths can be tuned without having to re-run the detector.

Volumetric Image Extraction

A volumetric image is generated for each detected object using the same beamforming algorithmic approach as for the target strength extraction. A cubic volumetric detection grid with an edge length of 1m and centered at the best focus location is constructed using a fine grid resolution of 5cm. The integrated target strength at each grid point is computed and stored to be later interpreted by the human expert in drawing conclusions on target geometry, burial state and burial orientation.

The software tool that we developed allows one to set certain parameters when submitting the post-processed scattering data to the trained RVM classifier. In the down-look case, one is a routine that can eliminate proud target detections. This is enabled by mapping the target scattering level versus lateral range which is effectively scattering level versus incidence angle at the water-sediment interface. (See Fig. 5.) These plots display the scattering level integrated over angle and frequency for the various paths on which the target was detected. These plots yield only the general fall off with range since different paths might access different target aspects e.g. beam, end, or quartering aspects. For a buried target, this curve falls to low levels due to lack of sediment penetration in the vicinity of the critical angle which for our sediment is about twenty-seven degrees. At the four meter down-look sonar altitude, this corresponds to a lateral range of about 7.9m. For proud targets, scattering levels remain high well beyond this range decreasing only due to spreading loss. One is also able to select the range over which the target scattering data is considered. Typically, we chose the range 3.5m to 9m. The lower range was selected in order to be cautious not to include scattering contributions from the water-sediment interface which can dominate the returns at short lateral ranges (near normal incidence). The upper range helped eliminate the low signal-to-noise data typically seen at the longer ranges for UXO-level returns.

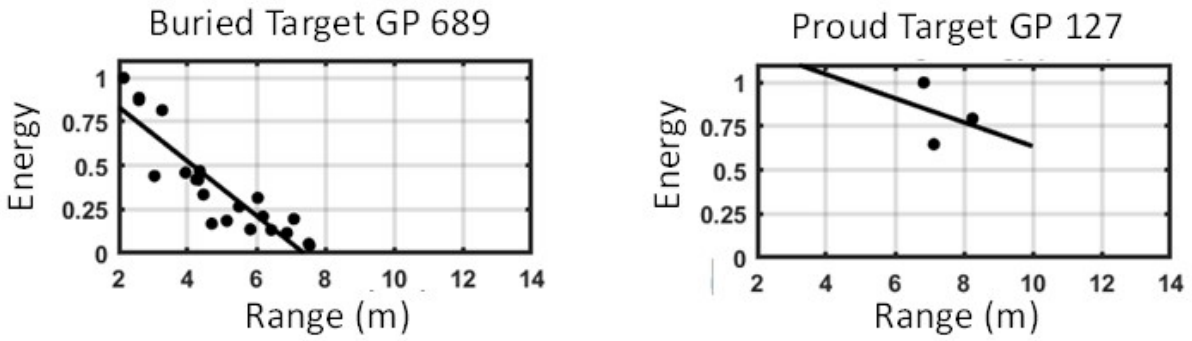

Figure 5. Plots of echo level versus lateral range obtained from the individual flight paths for a buried UXO (left) and a proud unknown target (right).

As an example taken from our ESTCP Blind Test Demonstration Program, we show in Fig. 6 the above-threshold display map generated after setting the parameters listed above and after the detections within 1.25m of one another have been aggregated. The global position identifiers (GP) are listed alongside each detection (circles). The detection threshold is set by extrapolation of the training field detection density to the corresponding area of the buried test field. In the example case here, the detection threshold was set by using the observed buried training field detection density number (0.1 detections/m^2) and extrapolating to the 10^4 m^2 area of the test field. In this case there are some 720 detections of interest presented in this display.

For the side-look data, post processing is carried out by beamforming the signals received by the 64 element hydrophone planar array onto the sediment surface of the proud target training field.

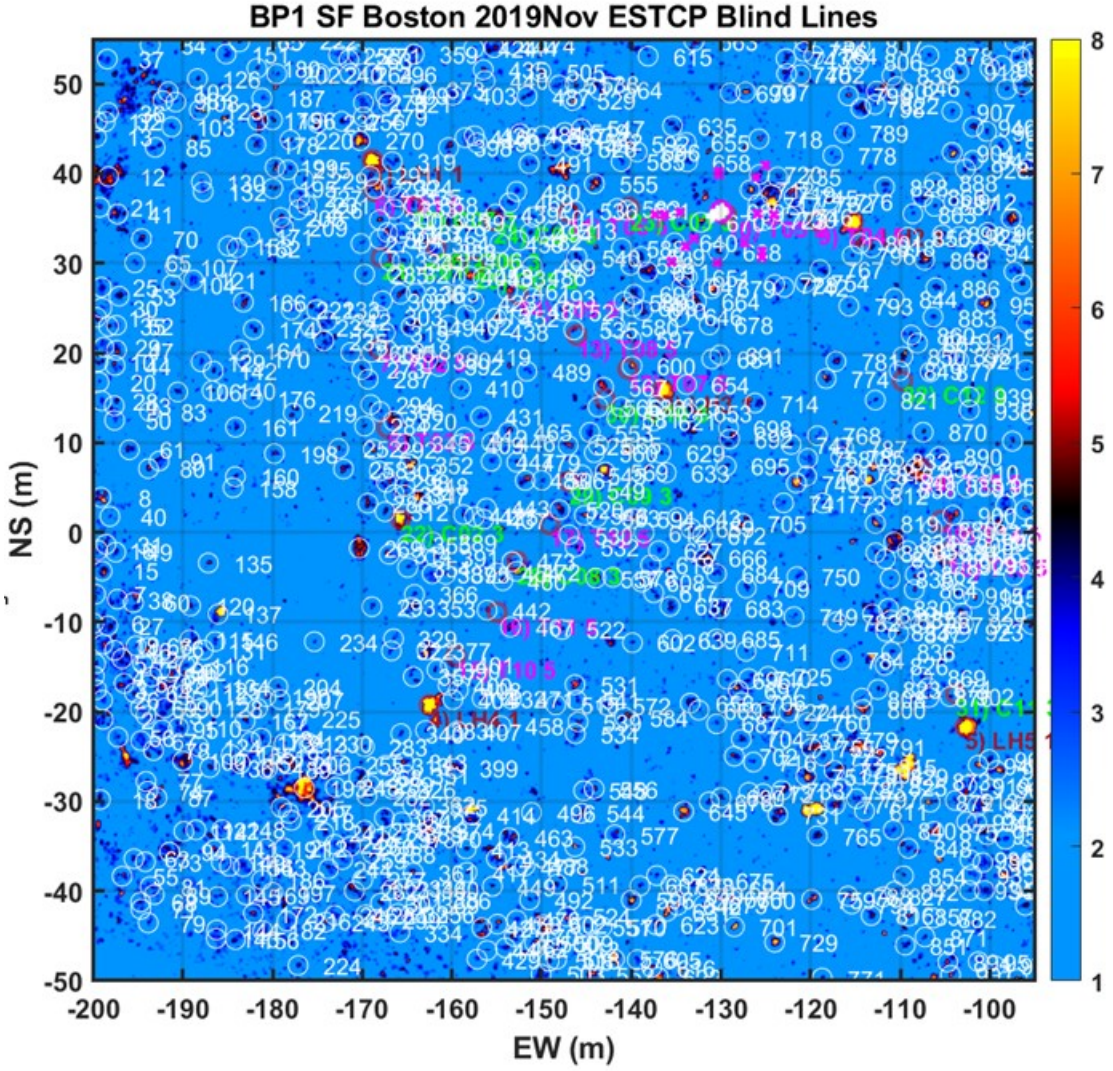

Figure 6. Detection display map showing the above-threshold detections obtained from the approximately 160 north/south, east/west, and diagonal flight paths in the test field of our ESTCP Demonstration Program. Each detection is labeled by its global position coordinate and the five navigation monuments are shown within the red circles.

The software tool that we developed also allows one to set certain parameters when submitting the post-processed side-look scattering data to the trained RVM classifier. For the side-look sonar, the major one allows us to select the lateral range over which the data is considered. In particular, the maximum range limit is determined by the UXO target strength levels and spreading loss which are parameters which for the most part control the signal-to-noise levels of the detected echoes. The minimum range limit set to prevent penetration into the sediment is determined by the acoustic projector beam width which allows acoustic energy to penetrate the sediment even when the

projector is aimed at the sediment surface at an angle below the critical angle. Initially, detection surfaces are produced for the data from each individual flight path. One then uses the unmistakable acoustic color patterns of the five Lincoln hat monuments to identify and locate them. The positions of the Lincoln hat monuments are then used iteratively to correct and align all the single path detections. From this, the global detection map is generated. This post-processed data is subsequently analyzed using a program we assembled that provides both acoustic color maps (frequency/angle plots) and time/angle plots for any desired detection location.

(G) Generation of Classification Call Lists

There are four features and three different pairs of feature parameters that we chose in our ESTCP Demonstration Program for the trained 2-feature RVM algorithms. The features are: (A) time correlation; (B) beam highlight width; (C) acoustic color symmetry correlation; and (D) detection map correlation. The three 2-feature sets for which we had trained the RVM algorithm are: A-B; D-A; and C-A. When the various features are extracted from the post-processed data at each of the above-threshold detections, they are applied to the trained algorithms and a corresponding call list is generated where the calls for the detections are ranked in the order of decreasing probability of being a UXO. One can also generate an "expert" call list wherein an expert familiar with the structural acoustics of UXO-like targets uses acoustic color and time-angle templates generated for the UXO targets from the training data, and even from other measurement and numerical studies. In the case of the down-look system, the target images could also be used to confirm various target calls.

(H) Construction of Final "Dig List"

As described in the above buried test field example, the two-feature RVM call list in an actual test field will contain some number of target detections, N_D, which in general will be larger (perhaps much larger) than the true number of UXO targets, N_{UXO}, present in the field. These detections are ranked by the RVM in the order of their relative probabilities of being a UXO with no cutoff threshold identified. To truncate the call list in an algorithmic manner in order to generate a smaller

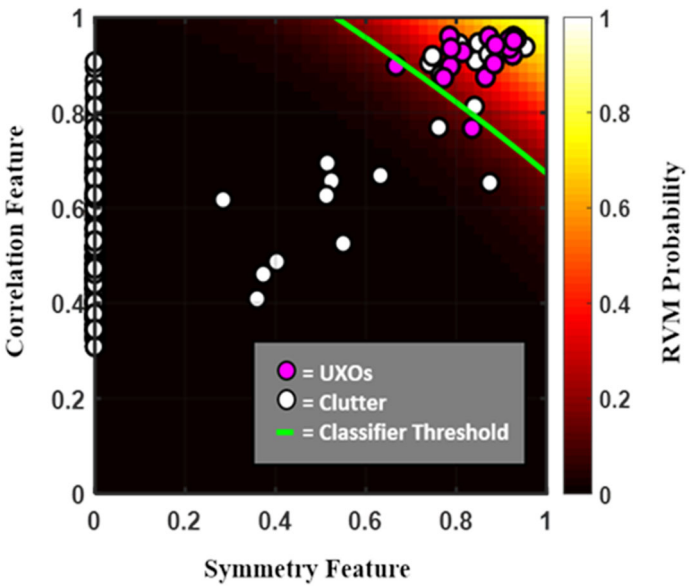

Figure 7. Feature space plots for detections in the proud blind test field overlaid with probability heat map from trained RVM model. Thresholding (green line) is shown using results from training exercise.

"dig list", a classification probability threshold could be derived *from the training data* by determining the maximum probability above which all UXOs are successfully called UXO in the *training* exercise and drawing the resulting contour line in the feature space plot. This is accomplished by "testing" each individual detection in the training field, extracting its corresponding features, applying them to the trained classification algorithm which used all the detected targets, and determining its ranked probability of being a UXO. To be able to do this successfully requires having sufficient statistical sampling in the training data. We believe we had sufficient statistical sampling in the proud test to be able to generate an effective classification threshold by this method. The contour line determined in this way for the NRL ESTCP Demonstration Program for the proud target field is shown in green in Fig.7 as applied to the proud *blind* test data. We display in Fig. 7 on the probability color map the first thirty-five calls from the call list which account for all fifteen UXO blind proud test targets. The threshold determined in this way from the proud training data shown by the green contour line in Fig. 7 (probability of 0.25) can be taken as the threshold for generating the smaller "dig list" from the test field ~ 300 detection call list. In this particular example, the "dig list" contains twenty-eight calls, and as can be seen from the figure, fourteen of the fifteen proud UXO fall on or above the threshold as do thirteen false targets. Regarding the "dig list," we would then have a $P_C = 0.93$ and a $P_{FA} = 0.46$. In other words, were this a clean-up operation, in prosecuting only these twenty-eight global detections at their global positions we would encounter fourteen of the fifteen UXO while encountering only thirteen false targets.

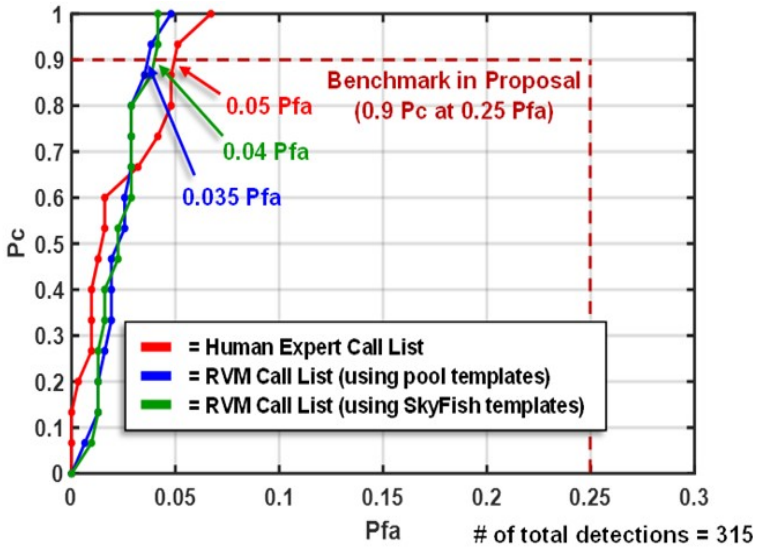

Figure 8. *The ROC curves associated with three call lists generated in the NRL ESTCP Demonstration Program for the proud blind test field. (Blue) RVM trained with free-field pool data; (Green) RVM trained using Down- look sonar buried target training data; (Red) Expert calls made using training templates and other available data.*

To put this in perspective, we show ROC curves generated in the blind proud test in the NRL Demonstration Program in Fig. 8 using ground truth information and the 315 member call list originally submitted "blindly" to the Program Office. One can see that the RVM generated ROC curves basically go through the point $P_C = 0.9$ and $P_{FA} = 0.04$. The latter number gives the number of false calls within the probability threshold determined above as 315 x 0.04 = 12.6 in agreement with the statement above *i.e. while encountering only thirteen false targets.* To emphasize by repetition: We can reduce the RVM call list of ~300 entries to a much smaller "dig list" of 28 detections to be prosecuted.

In contrast to the proud blind field case, we have not been as successful in generating a manageable "dig list" as was done in the proud case demonstration. At the present time, we do not believe we

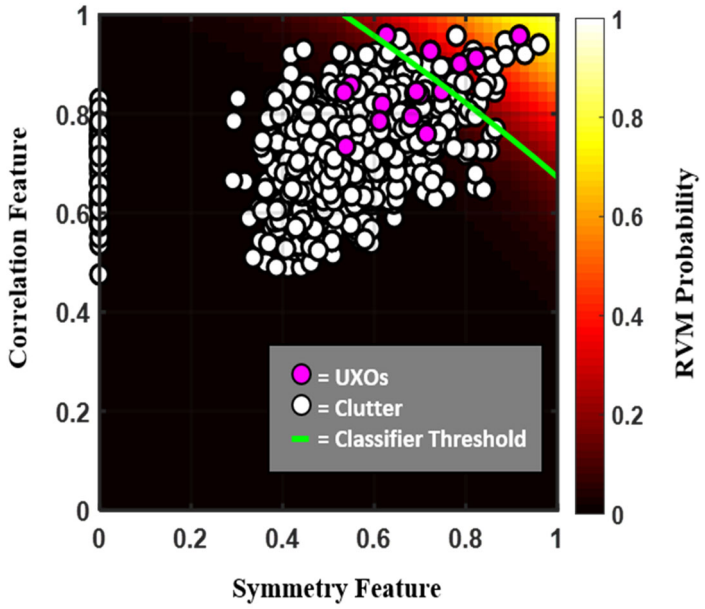

Figure 9. Feature space plots for detections in buried blind test field overlaid with probability heat map from trained RVM model. Threshold (green line) is shown using results from training exercise which was performed with a smaller, more limited data set compared to the testing field.

had taken sufficient buried target training data in the Demonstration Program to obtain an effective classification probability threshold and thus a compact buried target "dig list" for the buried target test. Our call list in the buried test field contains some 720 target detections. The contour line determined in the same manner as was done for the proud case is shown in green on the buried test data feature space plot in Fig. 9. As can be seen, there are many UXO detections below the green probability contour line as well as many false targets above it. We have determined that this poor threshold-setting performance is due to overtraining resulting from the comparatively limited sampling in the training exercise. In order to establish an effective means for setting the classification probability threshold for our RVM classifier in this manner, we would need to have collected more training data than was actually the case. The buried testing field was surveyed over a much larger area and thus much larger sediment volume resulting in a broader range of sediment variations and indigenous clutter. Basically, we needed to have increased the number of buried UXO training targets beyond the twelve that we used to perhaps fifty or so as estimated below. As a result of the limited statistics, the threshold determined using the buried training field feature space plot (see Fig. 9) correctly labeled only 5 of the 15 UXO targets while incurring 32 false calls. This poor threshold-setting performance for the buried targets due to overtraining is in contrast to the more successful situation experienced in the proud test which among other things did not involve sampling the sediment volume.

At the present time, we do not have a rigorous method allowing one to determine a-priori whether there is a sufficient statistical sampling in the training data to allow generation by the above method of a RVM probability threshold for obtaining an effective "dig list." However, we do have rough "rules of thumb" described in the following way. We believe that the number of SEEDS needed for training depends for the most part on three parameters: (1) the number of expected UXO types; (2) the range of UXO vertical burial angles (pitch) expected in the test environment; and (3) the various sediment types existing in the test environment. Regarding (1), this criteria would simply be the number of classes of UXO articles itself. Regarding (2), we suggest the following rough rule-of-thumb. If the buried targets are expected to be present over a range of vertical burial angles $\Delta\varphi$, then one should sample at least every φ separated by 10 degrees. So the number of buried

training targets for (2) in this case would be $\Delta\varphi$ / 10 for each target class. If it is expected that the UXO would be found buried at pitch angles between 0 degrees and 40 degrees, then one would need four targets of each class. And finally for (3), at this point we can only suggest that one needs to sample the sediment over the range of sediment types existing in the testing environment that are different enough to significantly change the acoustic scattering characteristics. Then considering all three parameters above, we describe the following example. Assume there are four expected UXO classes (as in our particular demonstration), a range of burial angles from zero to forty degrees, and three significantly different sediment types. In this case one would need 4 x 4 x 3 or 48 UXO training targets. In our buried demonstration, we had only twelve UXO. After the fact, we concluded that in the buried demonstration, our training data base was under-sampled. In fact, we have reported in our briefings and in the soon-to-be-released Final Report that this under-sampling affected both the ROC curve performance (although still meeting the goal) and the so-called "dig list." In particular, although not a demonstration goal, this under-sampling led to a larger than desired dig list.

(I) Lessons Learned

The important lessons learned in the execution of this demonstration program are as follows: (1) As just discussed, more attention must be paid to having a sufficient amount of training data. Given that the buried target case involves a large sediment volume which can include a significant amount of acoustic parameter variations and clutter, this issue is especially important for buried targets. For example, in answer to a later comment, we discuss rough rules of thumb in determining how many UXO targets should be used in the training fields in order to sufficiently sample both the UXO types and pitch angles and the spatial variations in the sediment. (2) Originally, the plan was to use high dimensional features as was done in earlier SERDP projects. However, we found that to avoid mathematical instabilities which exasperate the effects of the complex environment in our demonstration site, it was important to use low dimensional feature spaces. The very high dimensional feature spaces have the problem described by the phrase "curse of dimensionality" wherein instabilities arise due to the disparity between the high-dimensional feature space and the low-dimensional measurement space. Accordingly in our blind-test demonstration we migrated to feature spaces with low dimensionality. (3) The major costs of our demonstration exercises were related to the AUV size. These costs would be significantly lower with the use of smaller AUVs, and we believe our structural acoustic sonars can be implemented on smaller vehicles without much loss in performance.

Regarding further sonar system details, these are covered in detail in the soon-to-be-released Final Report. For UXO remediation, the details of the sonar hardware would not have associated security issues. They would likely have some ITAR restrictions but can be implemented on non-navy platforms. The transition pathway would be directly through a commercial process similar to what exists for more conventional in-water side-scan and sub-bottom surveys. The first part involves a manufacturer who provides and supports the equipment as a commercial product. The second part involves a company or companies that provide commercial operation services for site cleanup. NRL carried out the demonstration program with support from General Dynamics (GD is formerly Bluefin Robotics). GD's support included launching and recovery of the AUVs, carrying out required repairs and modifications, and general involvement during the actual training and testing runs. Such support is also available to any commercial firms using these AUVs. The

AUVs can operate from piers standing over waters with sufficient depths (e.g. ten feet or greater). The major cost driver by far is the choice of autonomous vehicle. The demonstration utilized 21" diameter AUVs which were already owned by NRL. With easily realized changes, the sonar technology could also be implemented on smaller, cheaper AUVs such as the IVER 3 class which are now available. Regarding the creation of a commercial vendor-supported hardware-software product, what is recommended is a product refinement effort that takes the existing sonar hardware design developed by NRL and refines it to be integrated onto a smaller AUV class than that used in this ESTCP effort. To make the switch to a smaller, more economical AUV platform would require some simple sonar redesign. However, based on the detailed studies which have been carried out in execution of this demonstration program - with respect to sonar design, fabrication, and classification performance dependence on source/receiver parameters - there is sufficient technical information to allow the successful design and procurement of nearly equivalent performance AUV-based structural acoustic sonars compatible with these smaller commercially available AUVs. A white paper is being prepared for submission to ESTCP to carry out such an effort. We recommend a migration step of the sonar hardware to a man-portable IVER3 or similar vehicle. This will result in an AUV-based system much simpler and less costly to ship and use. Moreover, it will significantly reduce the procurement cost of the system with the processing component largely unchanged from that used in the demonstration.