COMBINING MODEL-BASED AND *IN SITU* PERFORMANCE PREDICTION TO EVALUATE DETECTION & CLASSIFICATION PERFORMANCE

J. Gazagnaire Naval Surface Warfare Centre Panama City, Florida, USA
P. Beaujean Florida Atlantic University, Boca Raton, Florida, USA
J. Stack Naval Surface Warfare Centre Panama City, Florida, USA

1 INTRODUCTION

The mine countermeasures (MCM) problem typically involves applying search effort to gain either a desired percent clearance (*i.e.,* the number of mines found divided by the number initially present) or desired reduction in risk to transiting vessels. Planning for either of these two objectives requires an *a priori* estimate of search effectiveness. This estimate must account for both the probability of detecting and classifiying mine-like objects (MLOs), P_{dc}, and the probability of labeling a non-mine as an MLO (a false alarm), P_{fa}. The major contributing factors affecting P_{dc} and P_{fa} include not only predictable factors such as sensor performance and target characteristics but also variable environmental factors including sea state, currents, fine-scale bathymetry (*e.g.,* ripples, holes, *etc.*), sea floor characteristics (*e.g.,* sediment size, type, roughness, *etc.*), and the presence of clutter.

As many of these variable factors can not be forecast with sufficient accuracy, the *a priori* P_{dc} and P_{fa} estimates are inevitably inaccurate, and the resulting search effectiveness will not produce the desired clearance or risk. In instances where environmental conditions have deteriorated, a dramatic decrease in performance may result while unexpectedly favorable conditions may result in the over application of effort that could be better used elsewhere. Therefore, there is a substantial need to couple an *in situ* environmental characterization capability with a model-based performance prediction tool. This will facilitate not only an accurate assessment of the performance actually achieved but also the in-stride adaptation necessary to achieve the desired (planned) performance.

2 PREVIOUS & COMPLIMENTARY APPROACHES

This section briefly reviews other existing approaches to *in situ* performance estimation for the MCM problem. Most of these techniques are targeted toward sidescan imaging sonars as sidescan sonar is the primary MCM sensor. The major differences between the approaches are accuracy versus real-time operation and requirements on sonar-specific tuning versus operation given only generic sonar parameters.

A general yet relatively computationally expensive approach was developed at the NATO Naval Undersea Research Centre[1]. In this technique targets are inserted into actual sidescan sonar imagery to measure automated target recognition (ATR) algorithm performance in local environmental conditions. An iterative inversion process is used to extract bathymetry, reflectivity and beam pattern (sonar characteristics) from each image. A synthetic image is generated from the estimated parameters at each step in the iteration and compared to the original sonar image. The solution parameters are refined and the process is repeated until the pixel value error between the synthetic and actual image is sufficiently small. Bathymetry and reflectivity maps are then modified for the inserted target, allowing interaction with the 3-dimensional structure of the seafloor and accounting for target occlusion from taller objects that are closer to the sonar than the inserted targets. These simulated images are then tested against an ATR algorithm to measure performance in the collected sonar imagery.

An additional seafloor classification module segments the seafloor by texture; flat sand, sand ripples and complex, which includes rocks, seaweed, and sea grass. Clutter density maps are generated

and fused with the seabed classification maps. The model outputs are combined with the ATR outputs to generate P_{dc} and P_{fa} values by seafloor classification label.

A more computationally tractable method was developed with the aim of real-time in-vehicle mission evaluation. Espresso[2], a sonar modeling tool, is used to predict shadow and echo to mean background contrast levels for a specific target as a function of range. The contrast ratios are then used to scale the actual mean grey levels of the image, which are also a function of range. The result is a two-dimension mean shadow level and mean echo level versus range.

Target shapes are generated using mesh models, and the target is placed in scene. Ray tracing is used to determine the amplitude of the image at the location of the target. The ray tracing method does not take into account the interaction with the 3-dimensional landscape such as shadowing by bathymetry and distortion by sand ripples. The directionality of the beam is taken into account. The model uses probabilistic methods to predict the proportions of the shadow, background, and echo levels for each pixel that would appear at the target location. This probability map is used with the Espresso output described above to predict shadow and echo levels. The ATR algorithms are then run on the image data to estimate performance. The outputs of the model are receiver operator characteristic (ROC) curves and P_{dc} versus range.

Sensor specific efforts include an evaluation tool developed by the US Navy Research Lab, the Applied Research Lab—University of Texas (ARL-UT), and the US Naval Oceanographic Office for the US Navy's AN/AQS-20 sonar. This capability was demonstrated in December 2004[3]. The approach is to extract bathymetry and sediment information from the image. This information is then used to determine the doctrinal bottom type for the area, which then indicates the expected performance of the particular system and threat.

The AQS-20 system has both sidescan and volume sonars. The data from the sidescan is used to produce the clutter category and roughness information and the volume sonar is used to determine multibeam bathymetry and bottom impedance. The impedance is extracted from the data using the Acoustic Sediment Classification System embedded in SeaBED[4]. Empirical relationships are used to convert impedance to general bottom type (mud, sand and rock) and mine burial potential.

The US Office of Naval Research and ARL-UT have developed the Sonar Performance Monitoring System (SPMS) for the US Navy's AN/AQS-32 forward look sonar[5]. The SPMS generates, in real-time, a coverage map of the searched area. The reverberation level is measured through the sensor and estimated target strength values are used with these measured reverberation values to map the signal to noise ratio (SNR) for the entire search area. This SNR can then be used to determine the probability of detection and subsequently the percent coverage. This method works well for the systems that are primarily searching for volume targets and are solving the signal detection problem. For high resolution systems searching for bottom targets, the problem quickly becomes more complex.

3 MODELING & ESTIMATION OF PERFORMANCE

The MCM ATR community typically estimates performance as P_{dc} rather than separating P_d and P_c. However, this research handles their estimation separately and combines them in a final stage. The following section discusses the concepts supporting the models for P_d and P_c and then details how each is estimated and combined to achieve an overall performance estimate.

3.1 Modelling Classification Performance, P_c

In general, *classification* performance is difficult to predict *a priori* for sonar (*i.e.,* it is difficult to construct a robust sensor-environment-object model). This is due to the fact that many of the relevant parameters are highly situation dependent and vary with environmental conditions (*e.g.,* object pose, aspect, and range, uniqueness of shape, distribution of object pixels versus background pixels, *etc.*). Therefore, this research proposes a comprehensive model whereby the

predictable sensor-object relationships are thoroughly characterized beforehand while the unpredictable environment or situation dependent components are estimated and incorporated *in situ*. This model is an extension of that proposed by Manning[6].

For the estimation of P_c, the goal is to equate any given sensor-environment-object instance to an effective SNR, given by (1). As SNR_{eff} attempts to account for the major factors affecting performance, it may be explicitly linked to a known, pre-derived ROC curve defining classification performance (*i.e.,* P_c versus P_{fa} where P_d is assumed equal to 1.0).

$$SNR_{eff} = SNR_{im} + 15log_{10}(\alpha) + 15log_{10}(\beta) - \delta \text{ dB} \tag{1}$$

In this equation, SNR_{im} is the range-dependent sonar SNR. It is estimated as the maximum object highlight pixel value divided by the mean background pixel value in the un-normalized image. This estimate is preferred over one based on a model or the sonar equation because in practice sonar systems clip the brightest highlight pixel values well below their theoretical maximum (*e.g.,* to 255 for an 8-bit dynamic range). This substantially affects SNR_{im} and is much more amenable to direct measurement.

The α parameter is a statistical index that is a function of object size and resolution. It is a measure of the maximum sonar cross section or the maximum object dimension in pixels, and it attempts to capture object shape information. This parameter is estimated by considering objects of interest uniformly distributed over range and aspect. The β parameter is also a statistical index and is a function of sonar resolution only. It is a measure of the mean number of object highlight pixels expected for a nominal target distributed uniformly over aspect and range.

The δ parameter is a discount factor obtained by first modeling the image background pixel amplitudes as k-distributed random variables. The k-distribution[7] is given in (2) where $U(x)$ is the unit step function, K_υ is a modified Bessel function of the second kind with order υ, and $\Gamma(\cdot)$ is the gamma function. The shape parameter, υ, is related to the skewness of the distribution, and specific values of υ lead to other common distributions (*e.g.,* $\upsilon \to 0$ tends to a lognormal distribution, υ = ½ is a Wiebull distribution, and υ > ~20 tends to a Rayleigh distribution). The scale parameter, a, which is related to the background mean may be estimated as a function of the arithmetic sample mean, μ, and the shape parameter υ as demonstrated by Raghavan[6] and given here in (3).

$$p(x) = \frac{2}{a\Gamma(\upsilon)}\left(\frac{x}{2a}\right)^{\upsilon} K_{\upsilon-1}\left(\frac{x}{a}\right)U(x); \quad \upsilon > 0 \tag{2}$$

where

$$a = \frac{\mu\,\Gamma(\upsilon)}{2\Gamma(\upsilon + \tfrac{1}{2})\Gamma(1.5)} \tag{3}$$

The proficiency of the k-distribution has been demonstrated in both radar and sonar[6,9]. The motivation for estimating δ via this distribution may be seen by examining the k-distribution behaviour over multiple values of υ as illustrated in Figure 1. If a given sonar image has background pixels that are Rayleigh distributed (large υ), then there should be little overlap between these background pixels and either the shadow pixels (small amplitudes) or highlight pixels (large amplitudes). In this instance, shadow, highlight, and background pixels are easily separated; therefore, the classification problem is simplified and there should be little or no discount applied to SNR_{eff} as estimated by the first three terms of (1).

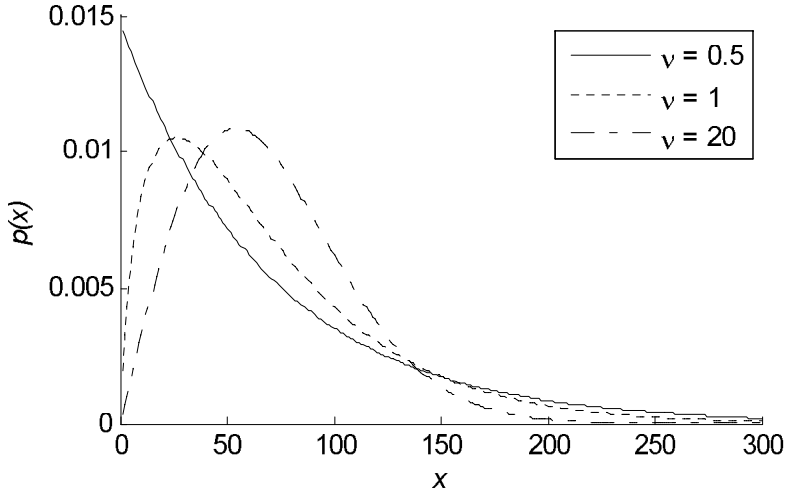

Figure 1. K-distribution for several values of the shape parameter. Note small values of υ produce a heavy-tailed distribution while large values produce a more concentrated distribution.

When a sonar image background is lognormal distributed ($\upsilon \to 0$) the background pixels will substantially overlap the shadow pixels as $p(x)$ (*i.e.,* the background pixel distribution) is seen to increase as $x \to 0$. The background pixels will also overlap the highlight pixels as $p(x)$ is also seen to have a heavier tail for $\upsilon \to 0$. Therefore, the classification performance is confounded by the additional difficultly in separating object pixels from the background pixels, and a commensurate reduction should be applied to SNR_{eff}. Figure 2 illustrates typical seafloor bottom types and their corresponding distributions for reference.

3.2 Modelling Detection Performance, *P*d

Unlike classical problems from detection theory that often involve sophisticated statistical tests to pull signals out of noise, *detection* of MLOs in sonar imagery is typically not a SNR-limited problem. This is because imaging sonar operation must be restricted to a range containing enough SNR (and spatial resolution, signal information content, *etc.*) to resolve an object and differentiate it from the myriad of mine-like non-targets littering the ocean. This restricted operating range typically results in an SNR that is more than sufficient to perform detection.

Two of the main factors that do affect MLO detection are reverberation and objects not being in the field of view (sometimes noted as percent hidden). The contributing factors to percent hidden are burial and occlusion. Estimating burial potential is an active research topic where various groups are investigating impact burial as well as scour from hydrodynamic and sediment properties. A brief review of the art is presented by Rennie and Brandt[10]. As burial estimation requires knowledge of parameters that can not be inferred from imagery alone, this research considers percent burial as an input provided *a priori* from some other technique.

Occlusion occurs when the shadow of one object is cast over another object preventing its insonification. The average likelihood of occlusion can be captured by a measure of bottom roughness. The other main factor, reverberation, is an undesired effect resulting in energy scattered off the surface and bottom. When there are a large number of significant bottom scatterers (bottom reverberation) or significant surface wave effects (surface reverberation), the result is often a rise in average pixel values to wash out highlights or a rise in minimum pixel value to wash out shadows respectively. In extreme cases, the entire image content may be washed out. In this research, an initial P_d of 1.0 is assumed and reduced appropriately based on estimates of occlusion and shadow contrast, which is used as an estimate of reverberation level.

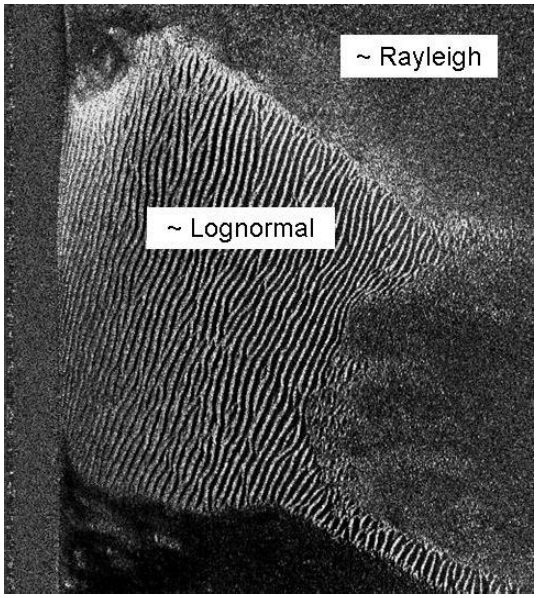

Figure 2. Example backgrounds in a typical sonar image. The top / right is flat smooth sand (Rayleigh), the middle is rippled sand (lognormal), while the bottom is mud (almost all pixels have zero amplitude).

3.3 Performance Estimation

This section explains how each of the aforementioned parameters are estimated (either via models or *in situ*) and how they are combined to achieve an overall performance measure. The process is depicted in Figure 3 where the output is expressed in terms of false alarms, *FA*, per area (sometimes expressed as false alarms per image) and $P(y)$ which is P_{dc} versus range.

In this figure, the two dashed boxes delineate which parameters are assumed to be fixed or specified versus the ones known to be variable. The variable parameters may then be either predicted via a model or estimated *in situ*. This model-based prediction capability is required for *a priori* planning purposes (*i.e.,* before any sensor data is collected). However, once sensor data is available, the process may switch to *in situ* estimation and realize more accurate and *local* estimations of performance.

To compute α and β in this model, a synthetic dataset of statistically substantial size is utilized over a wide variety of target types, aspects, burials, and ranges, sonar SNR and resolution, and background distributions. This dataset is an extension of that developed and described by Manning[6]. From this data, a mapping from object size and resolution to α is learned by computing the expectation of the maximum object dimension in pixels over range and aspect. A mapping from resolution to β is learned by computing the expectation of the total number of object highlight pixels for a nominal spherical target distributed uniformly over range. The minimum target range input indicates how much of the nadir is ignored per image while the sonar altitude affects the effective resolution at range.

To compute SNR_{eff}, the α and β parameters must be combined with SNR_{img} and δ. These additional parameters may be either estimated from a model or from imagery as illustrated in Figure 3. To estimate δ from imagery, an approximately 1.7 m by 1.7 m window is moved across each image and υ is computed via Raghavan's method[6] for each window. The synthetic dataset is then used to learn a mapping between δ and υ by measuring the decrease in P_c as υ is varied from approximately 25 to 0. δ is then chosen to apply the appropriate discount to SNR_{eff} (and commensurately P_c) as υ decreases. The imagery based estimate of SNR_{img} is again simply

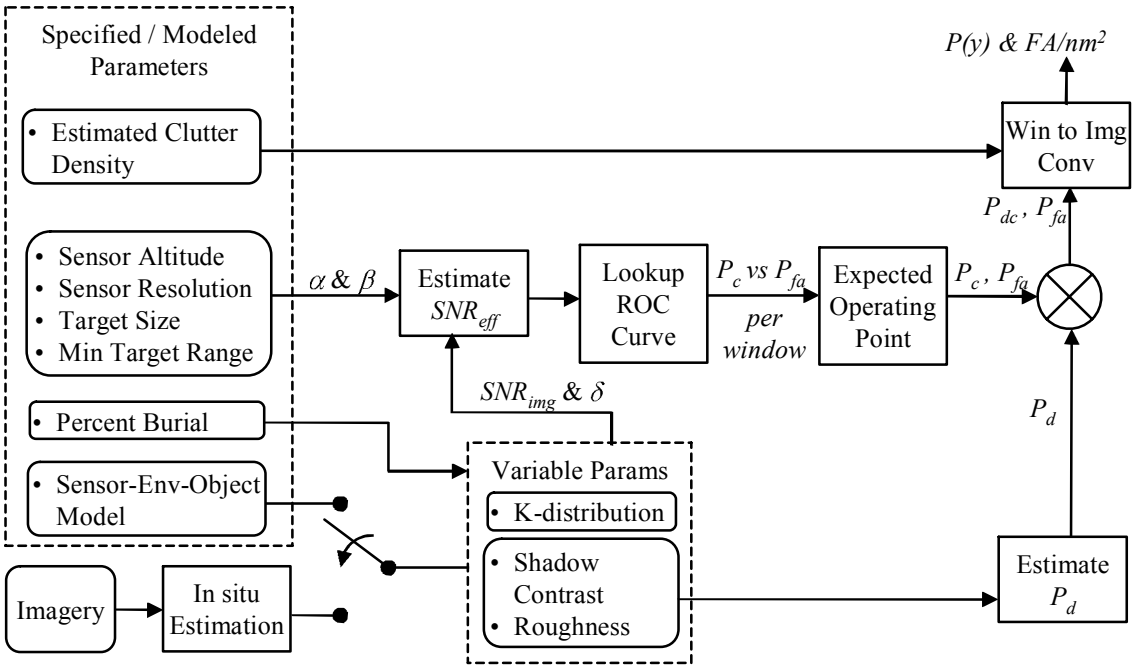

Figure 3. Flow chart for the proposed performance prediction capability. Dashed boxes note specified vs. variable parameters where the variable parameters may be either modeled or estimated *in situ*.

the maximum pixel value divided by mean pixel value in the window. These four parameters are then combined via (1) to produce SNR_{eff} per window. This process is performed for all windows across the image (applied with approximately a 66% overlap). The synthetic data is also used to learn a ROC curve (*i.e.,* P_c versus P_{fa} with P_d = 1.0) as a function of only SNR_{eff}. Thus any value of SNR_{eff} may be used to select the appropriate ROC curve from the suite of pre-computed ROC curves.

For the model-based estimate of SNR_{img} and δ, multiple sonar types and models are constructed *a priori* and made available to the user. These include models for well-known, well-characterized sonars as well as generic sidescan and synthetic aperture sonars. In the model-based approach, the user specifies the sonar type and bottom roughness. From this, the k-distribution (υ) , shadow contrast and SNR_{img} are produced while δ is again inferred from υ. The model used for this purpose is PC-SWAT[11].

For the *in situ* estimation of roughness, a 2-dimensional autocorrelation is computed for each image in an attempt to equate roughness as a function of correlation. The principle idea behind the approach is that more correlated bottoms such as sand ripples and groups of large objects will have a larger roughness parameter using this technique. Roughness parameters are estimated as a function of the image autocorrelation coefficient values.

The *in situ* estimation of shadow contrast is based on an estimate of the noise floor across all pixels of a fixed range. In general, the noise floor will raise as a function of range and this has the effect or raising the mean shadow pixel value to that of the background. Thus shadow contrast is estimated as the difference between the background level and the noise floor. Once υ, shadow contrast, and roughness are estimated (by either method), a direct mapping is applied to get P_d. This mapping is again learned via synthetic data and applied as a function of only these three parameters.

Referring again to Figure 3, note the output of the ROC curve lookup is P_c versus P_{fa} for each window. An expected operating is chosen as P_{fa} = 0.1 for generic sonars and tuned around this

value for specific sensors. From this value, the corresponding P_c is multiplied by P_d to produce a P_{dc} for that window. In the final stage, an estimate of clutter density is injected to convert P_{fa} to *FA* per area. The P_{dc} estimates per window are then concatenated in range to form $P(y)$ curves every 1.7 m in cross range. this result is also averaged across cross-range to produce an aggregate $P(y)$ per image.

4 PERFORMANCE PREDICTION TOOL

This capability is integrated into a $P(y)$ Performance Prediction Model / Simulation application whose graphical user interface (GUI) is illustrated in Figure 4. In this GUI, *Select Run Option* is used to select between a PC-SWAT-based model or raw sonar imagery as indicated in Figure 3. Within Figure 4, the left plot shows P_{dc} and P_{fa} versus range while some of the key estimated parameters are shown versus range in the right plot.

The performance of this tool has been informally validated for several US Navy experimental systems including the Marine Sonic 900 and Klein 5000 sidescan sonars. It has also been used for the US Navy Fleet systems AN/AQS-20 and AN/AQS-14 resulting in its selection by the US Navy Expeditionary Warfare Division as the "trusted performance prediction model" for defining system capabilities.

5 CONCLUSIONS

In conclusion, Manning[6] has previously demonstrated that classification performance can be accurately predicted by SNR_{eff} in Rayleigh-distributed backgrounds. The work presented herein extends this result by accounting for a wide variety of background statistics as captured by the shape parameter of the k-distribution. It also presents a real time tool (or offline simulation) capable of accurate performance prediction using either models and *a prior* specifications or actual imagery gathered *in situ*. The extremely low computational complexity and cost of this technique stems from the fact that the mapping between the key parameters and performance is learned offline and stored for use in real time. Thus, these key parameters may be estimated *in situ* (or derived from a model) and immediately combined to produce an accurate performance estimate.

The future research of this work concern more robust estimation techniques for shadow contrast and roughness. Although not discussed, the primary method for ascribing roughness to an area is via (US) Naval Warfare Publication NWP 3-15.41. While this method is used because it is US Navy Doctrine, higher fidelity techniques are under investigation.

6 REFERENCES

1. Y. Petillot, S. Reed, and V. Myers, "Mission planning and evaluation minehunting AUVs with sidescan sonar: Mixing real and simulated data", Tech. Rpt. SR-447, NATO Undersea Research Centre, La Spezia, Italy, Dec. 2005.
2. G. Davies, "ESPRESSO scientific user's manual," Tech. Rpt. M-154, NATO Undersea Research Centre, La Spezia, Italy, Dec. 2005.
3. M.M. Harris, W.E. Avera, C.A. Steed, and J. Sample, "Through-the-sensor determination of AN/AQS-20 sensor performance Demonstration 1, December 13 through 17," Tech. Rpt. US Naval Research Laboratory NRL/FR/7440-05-10,106, Washington, DC, Jun. 2005.
4. D. N. Lambert, D. J. Walter, D.C. Young, M. D. Richardson, "High resolution acoustic seafloor classification system for mine countermeasures operations," *Int. Conf. Underwater Acoustics*, Univ. New South Wales, Dec. 5-7, 1994.
5. N.P. Chotiros, J.M. Estes, K.N. Scarborough, S.P.Pitt, and R.V. Klinksiek, "Sonar coverage mapping," *Society for Counter Ordnance Tech. 5th Int. Symp. Technology & Mine Problem,* http://www.demine.org, Apr. 22-25, Monterey, CA, 2002.

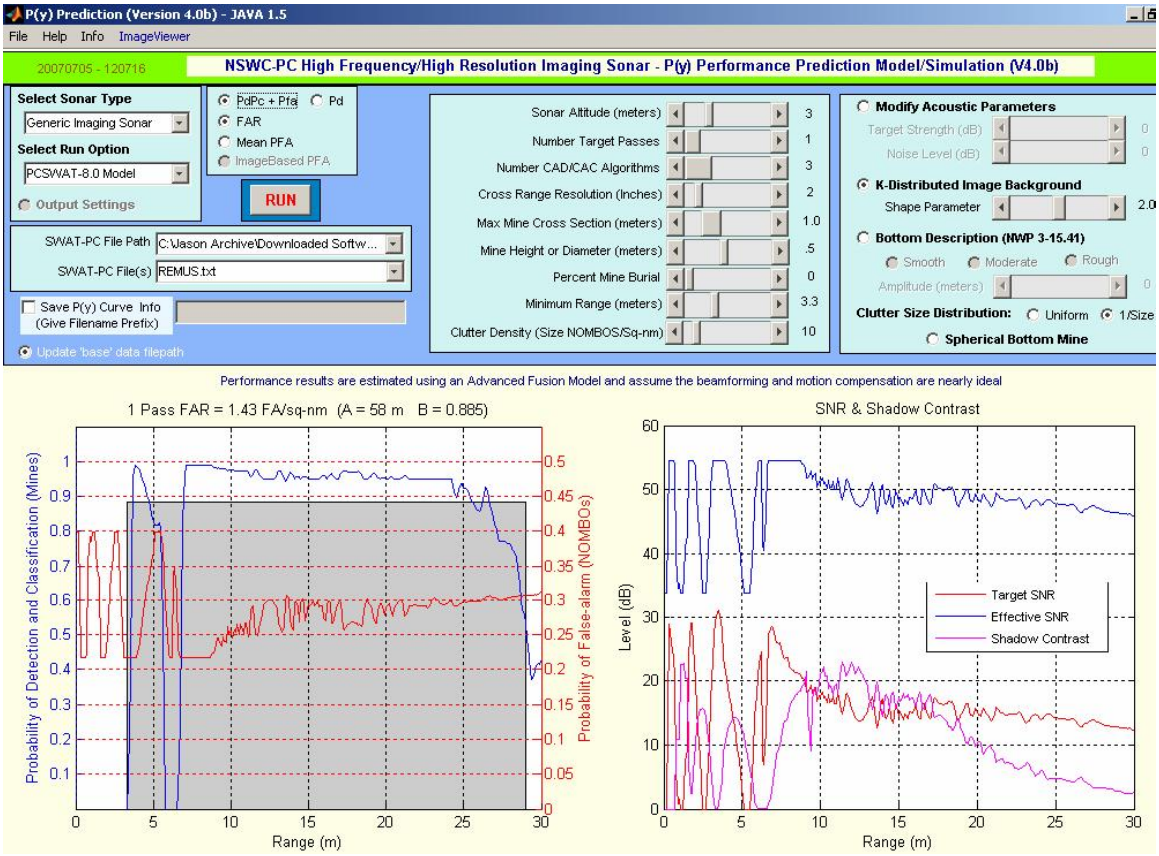

Figure 4. Graphical user interface for the *P*(*y*) Performance Prediction Model / Simulation application. This application may use PC-SWAT models or raw imagery as inputs. It may also run in real time onboard a vehicle without the user interface.

6. R. Manning, "Small object classification performance of high-resolution imaging sonars as a function of image resolution," *IEEE/MTS OCEANS,* vol. 4, pp. 2156-2163, 2002.
7. E. Jakeman and R.J.A. Tough, "Generalized K distribution: A statistical model for weak scattering," *J. Opt. Soc. Amer.*, vol. 4, no. 9, pp. 1764-1772, Sept. 1987.
8. R. S. Raghavan, "A method for estimating parameters of k-distributed clutter, *IEEE Trans. Aero. Elec. Sys.,* vol. 27, no. 2, Mar. 1992.
9. J. Dunlop, "Statistical modelling of sidescan sonar images," *IEEE/MTS OCEANS*, vol. 1, pp. 33-38, 1997.
10. S.E. Rennie and A. Brandt, "A probabilistic expert system approach for sear mine burial prediction," *Society for Counter Ordnance Tech. 5th Int. Symp. Technology & Mine Problem,* http://www.demine.org, Apr. 22-25, Monterey, CA, 2002.
11. G.S. Sammelmann, "Propagation and scattering in very shallow water," *IEEE/MTS OCEANS*, vol. 1, pp. 337.344, 2001.