

Quantifying the Effects of Propagation on Classification of Cetacean Vocalizations

Paul C. Hines
Department of Electrical and Computer Engineering
5269 Morris St.
Murray Building, Room 200
PO Box 15000
Halifax, NS Canada B3H 4R2
phone: (902) 809-0559-3100 email: phines50@gmail.com

Carolyn M Binder
Department of Oceanography, Room 5638
Dalhousie University
Halifax, NS Canada B3H 4R2

Award Number: N000141410237
<http://www.dal.ca/faculty/engineering.html>

LONG-TERM GOALS

To develop a robust automatic classifier with a high probability of detection and a low false alarm rate that can classify vocalizations from a variety of cetacean species in diverse ocean environments.

OBJECTIVES

In previous work as part of ONR grant N000141210139 a unique automatic classifier developed by the PI that uses perceptual signal features –features similar to those employed by the human auditory system– was employed to successfully classify anthropogenic transients, and vocalizations from five cetacean species. Although this is a significant achievement, successful implementation of this (or any) classifier requires that it be temporally and spatially robust. The primary goal will be to address the question: “Will it work on vocalization data from these species collected under different environmental conditions?” To examine this, discriminant analysis will be used to rank the aural features in terms of their ability to separate the vocalizations between species. Then, the more highly ranked features will be tested for robustness. This will be done by performing a propagation experiment using cetacean vocalizations and synthetically generated calls as source signals, and testing the received signals with the classifier. The measurements will be complemented by comparing experimental results to propagation model results with the goal of generalizing the results to other ocean environments.

APPROACH

The research is part of a PhD program undertaken by Ms. Carolyn Binder under the supervision of Dr. Paul C. Hines. The postgraduate program is being conducted collaboratively in the Oceanography and

Electrical Engineering departments at Dalhousie University where Dr. Hines holds adjunct professor and research posts, and at Defence R&D Canada–Atlantic where Ms. Binder is a researcher.

Passive acoustic monitoring (PAM) is widely in use to study marine mammals; since marine mammals can be found in all ocean basins, their habitats cover diverse underwater environments. It is well known that acoustic propagation can vary substantially between environments which can result in distortion of acoustic signals [1-3]. This in turn, can lead to environment-dependent time-frequency characteristics of a received vocalization. The resulting distortion of vocalizations may impact the accuracy of PAM systems. Thus, to develop a classification system capable of operating in many environments one must understand the role of propagation on the classifier.

A prototype aural classifier developed at Defence Research and Development Canada has successfully been used for inter-species discrimination of cetaceans [4]. The aural classifier is an effective PAM tool because it employs perceptual signal features, which model features used by the human auditory system [5]. The proposed research aims to examine the robustness of this classifier, and the perceptual features it uses, to environmental conditions. To accomplish this, propagation experiments were conducted by transmitting a set of real and synthetic vocalizations from bowhead and humpback whales, and measuring the received signals at a variety of ranges [6]. The transmitted and received signals will be tested using the aural classifier to identify any performance degradation due to propagation. The measurements will be complemented by propagation-model results using environmental inputs measured during the experiments. The model results will provide physical insight into what propagation effects have the greatest impact on the classifier's robustness, and aid in generalizing the experimental results to other ocean environments. It also enables one to put bounds on realistic within-environment variability. It is worthy of mention that there is no study published in the literature that systematically analyzes the impacts of propagation on an automated classifier, using both underwater propagation experiments and complementary modeling.

If it is found (as is anticipated) that propagation does impact some features, then the next step is to rank the features in order of their sensitivity to propagation-related effects. Features which are especially sensitive to the acoustic environment might simply be removed from the aural classifier. Alternatively, it may be found that many of the perceptual features are environment-sensitive and therefore it is unreasonable to exclude all of them. In this case, a strategy may instead be developed to generate training sets for the classifier that take propagation-related signal distortion into account; this could be done either by acoustically distorting the training data by propagating them through a modeled environment or by including vocalizations in the training set that came from a variety of propagation ranges as done in [7, 8].

WORK COMPLETED (FY 2014)

The focus of the effort during FY 2014 has been two-fold:

1. Identify a propagation model suitable for model-data comparison and perform preliminary modeling using that model.
2. Begin detailed analysis of the data collected during the propagation experiments to test the robustness of the aural classification feature set. Details of the experiment are reported in [6] and are not repeated here due to space limitations, but a schematic of the experimental geometry is shown in Figure 1.

Propagation model selection: A variety of numerical techniques have been developed to study ocean acoustic propagation. These include ray tracing, normal modes, parabolic equation, wavenumber integration, and finite element models [2, 9]; each has its advantages and no single method is suitable for handling all the possible environmental conditions, frequencies, and transmission ranges that are dictated by various applications [10]. There were several *essential* features as well as some additional *desired* features taken into account during model selection: The essential features of the model are:

- It must have an accurate representation of bottom interactions since much of the cetacean PAM occurs in shallow water [11].
- It must be capable of handling a wide frequency range to encompass the low frequency range of bowhead and humpback vocalizations (extending down to approximately 50 Hz) as well as the higher frequencies used for the propagation experiments (1–4 kHz)¹.
- It must be accurate over the relevant spatial scales (tens to hundreds of kilometers).
- It must be capable of performing pulse propagation that can simulate a time-domain representation of a signal after it has been propagated through a modeled environment.

In addition to these essential features, it is desirable if the model is:

- range-dependent since shallow water environments are often range-dependent,
- ‘public domain’ or non-proprietary to facilitate collaboration and ease of access,
- reasonably easy to implement.

The Ocean Acoustics and Seismic Exploration Synthesis (OASES) model, a general-purpose computer code for modeling seismoacoustic propagation was chosen [12]. It can be used to propagate signals through a user-defined environment using wavenumber integration which provides extremely accurate results; and it can determine the acoustic field for multiple sources and receivers simultaneously. OASES is a modern version of the SAFARI model and has the reputation of being very efficient code. The pulse propagation module in OASES known as OASP, is easily interfaced with all the other processing modules of OASES (e.g., transmission loss), and the range-independent modules were freely available for download. The model outputs the received pressure time series for each source signal. The time series is representative of a signal recorded on a hydrophone located at depth z in the water column after the signal was transmitted range R through the ocean environment. OASES-OASP has previously been used to localize North Atlantic right whales because of its ability to properly treat the physical effects typical of a shallow-water waveguide [13]. OASES-OASP can handle range-dependent environments if the appropriate modules are licensed and implemented, but the range-dependent modules have yet to be implemented in this study.

Although OASES was selected as the primary model for this research, it does not preclude the use of other models. For example, ray-based models may be employed during the course of the research to provide additional insight into the acoustic propagation physics. There is insufficient space to review

¹ All vocalizations –real and synthetic– were up-sampled to the 1–4 kHz frequency band prior to transmission. Although this was motivated by limitations in transducer bandwidth, transmitting signals in this higher band had several advantages: experiments were shorter in duration due to the time compression of the pulse sequence; high frequency signals attenuate more rapidly than low frequency signals so that the spatial extent of the experiments was decreased, thereby reducing the environmental impact; and the potential of behavioral impact on marine mammals in the area was reduced because the signals would not be misinterpreted as real whale vocalizations by marine life.

all the modeling approaches and computer implementations associated with them. Instead, the salient points associated with them are summarized in Table I, which has been inserted at the end of the report due to its size.

Experimental Data Analysis: The acoustic data that were collected during the experiments are currently being processed. The processing stream is as follows: The signals are identified in the recordings made by each hydrophone using the known time that signals were transmitted and correcting for the drift rate of the ship. After each signal has been detected, a four second segment of the signal is extracted with the detection located approximately in the centre of the segment. Each extracted detection is saved to a WAV file and high-pass filtered to remove the DC-offset applied by the recording equipment. At this point, the received signals are able to be input to the aural classification algorithm.

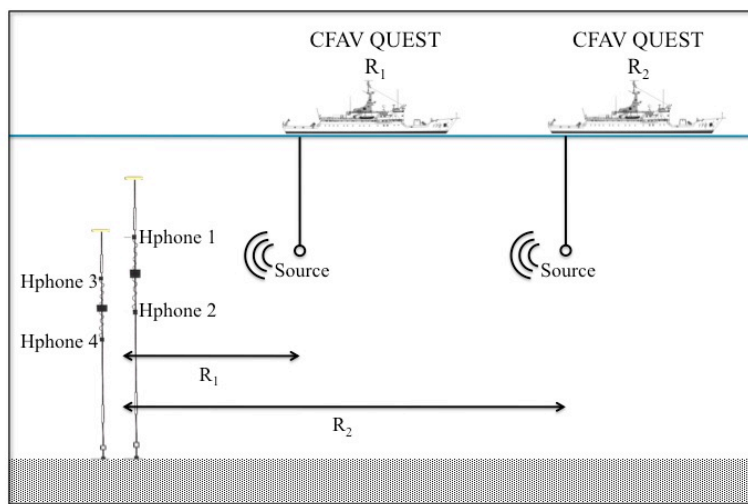


Figure 1: Representation of the experimental setup. The ship first deployed two hydrophone moorings, moved to the first location and transmitted the set of signals, then moved further away from the recorders and retransmitted the signals. R1 and R2 represent the horizontal range the signals propagated from the source to the midpoint between the moorings.

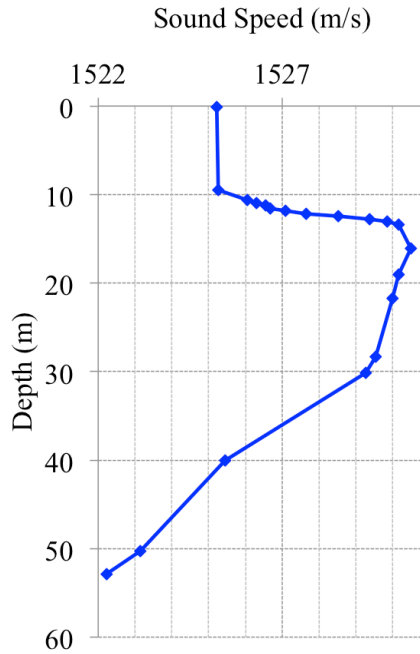


Figure 2: Sound speed profile measured during the propagation experiment that was used in OASES-OASP modeling.

The 58 perceptual features are calculated for each signal transmitted during the experiments and the signals received on each of the deployed hydrophones. An initial assessment of the classifier performance is accomplished by training the classifier on data recorded by a monitor hydrophone that was deployed from QUEST during the experiments, and testing on data transmitted through the water. Classification tasks are divided so that the real and synthetic whale calls will be considered separately.

Table II: The five features that best discriminate between example bowhead and humpback whale calls.

Feature	Description
<i>Duration</i>	Time delay between the start and end of the signal
<i>Global mean subband decay time</i>	Average time delay between the peak of the temporal envelope and the end of the signal
<i>Local maximum subband decay time</i>	Maximum time delay between the peak of the temporal envelope and end of the signal; each subband considered individually
<i>Frequency of global maximum subband attack time</i>	The center frequency of the filter bank channel which contains the maximum subband attack time (i.e., the time delay between the start of the signal and the peak of the temporal envelope)
<i>Peak loudness value</i>	Maximum value of the perceptual loudness spectrum.

The classifier is first trained using the five features that best discriminated between example bowhead and humpback vocalizations (refer to Table II). Examining and comparing the decision regions for signals transmitted over each of the ranges will allow qualitative analysis of the classifier’s robustness. Quantitative analysis will consist of comparing the accuracies and area under the ROC curve AUCs, as well as the class means and variances. If propagation affects the perceptual features used for

classification, then one would expect that the class means of the propagated signals in the training and testing sets would be significantly different, and/or there would be a large change in the variance of the classes. This will likely result in a decrease in the AUC and/or the classification accuracy. This type of analysis will be repeated allowing the classifier-algorithm freedom to select the best perceptual features to train the classifier.

RESULTS

Preliminary results from the aural classifier are presented in Figure 3 as performance matrices. These data are for the synthetic vocalizations and were collected with a hydrophone deployed at 29 m depth. The 0 km range refers to the original signals (real and synthetic) after up-sampling, but not subject to the signal conditioning designed to flatten the TVR curve of the transducer, nor transmitted into the water; the 70 m range refers to signals which were recorded on the monitor hydrophone deployed from QUEST’s well deck. Each element in these performance matrices represents the mean performance, obtained using 5-fold cross-validation, of the classifier using the five features that best discriminated between the real bowhead and humpback vocalizations. For example, the square in the fifth column and third row of the left hand panel of Figure 3 represents a mean AUC = 0.95 determined by training the aural classifier on signals transmitted over 10 km and then validating (or testing) the classifier on data transmitted over 1 km.

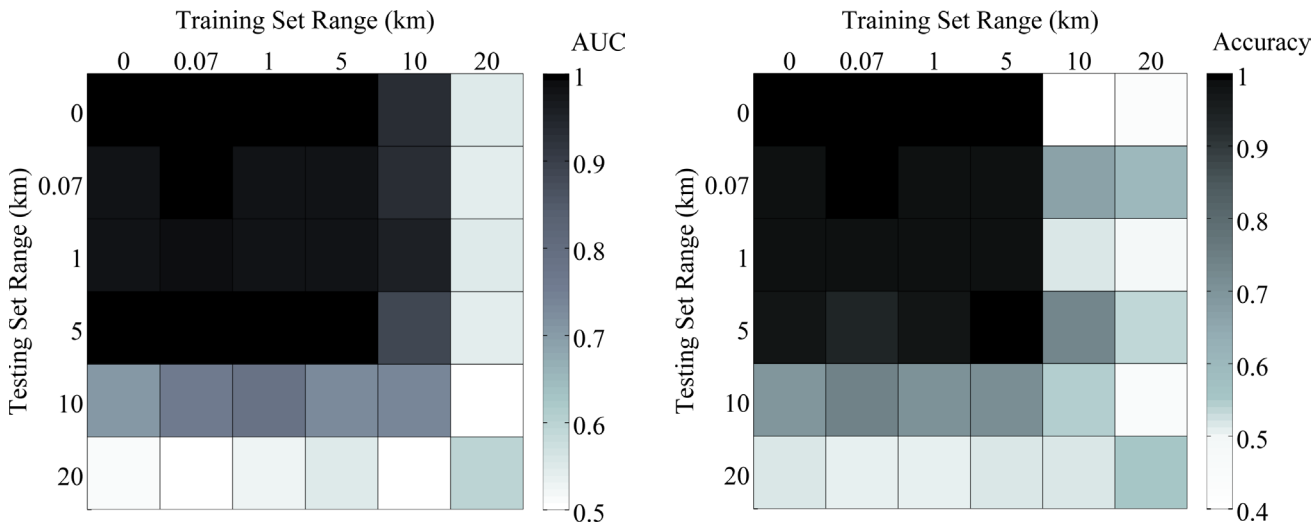


Figure 3: Aural classifier performance matrices generated from the sea trial data. These represent classification performance in terms of AUC (lhs) and classification accuracy (rhs) for all combinations of training the classifier on signals transmitted x km and testing on signals transmitted y km.

These preliminary results from the propagation experiments suggest that there is a range-dependent component to the classifier performance. The classifier performance for all training/testing pairs that included data transmitted over 20 km was little better than chance — likely because the SNR of the received signals was so low that when these ‘detections’ were listened to, no signal could be discerned from the ambient noise. Also of note is that the elements above the main diagonal are darker (i.e., represent better performance) than those below the main diagonal. A similar pattern was noted in the SNR-dependence investigation of Murphy and Hines [7], suggesting that SNR may, at least in part, be

driving the decrease in classifier performance. Mouy et al. [14] also noted that false negative rates increased as SNR decreased. Further work is required to determine if the performance degradation noted here is a result of complex propagation effects or a decrease in SNR as signals propagate over longer ranges.

The OASES-OASP model was run for geometries and environmental parameters consistent with the Gulf of Mexico sea trial reported on in [6]. The environment was modeled as a horizontally stratified water column overlaying a sediment half-space. The geoacoustic parameters of the sediment half-space (listed in Table III) were selected to be consistent with the silty-clay sediment type observed at the experimental site. The sound speed profile shown in Figure 2 was measured with a CTD during the experiment and was used to model the water column characteristics. The following results considered only the propagation of the *synthetic* signals with the source and receiver located at depths of 30 m and 36 m, respectively.

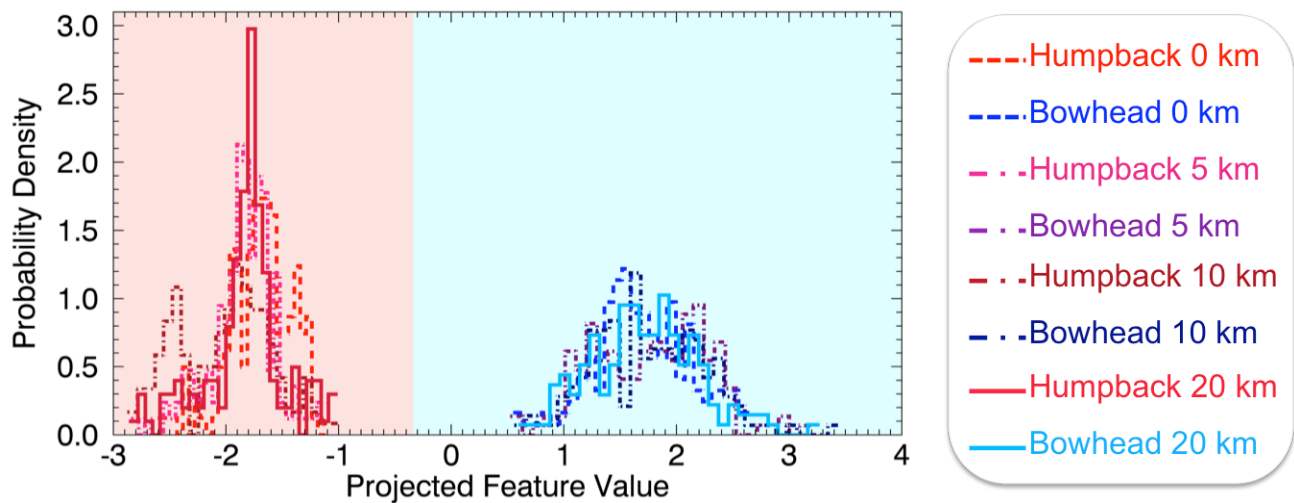


Figure 4: Aural classifications results for synthetic whale calls propagated through the OASES model over 0–20 km ranges. Data points have been binned into histograms after being projected onto the horizontal axis. The decision threshold (represented by the line separating blue and pink backgrounds) was determined by training the classifier on signals ‘propagated’ through the modeled environment over 0 km, then the classifier was validated on signals propagated over 5, 10, and 20 km ranges.

Figure 4 shows results obtained by training the classifier on synthetic signals propagated with OASES-OASP over a range of 0 km. The 0 km range is obtained by placing a receiver at the same range step as the source to capture any processing effects of the model. This plot shows histograms of projected feature values that were determined by performing linear discriminant analysis with the five features that best discriminated between the example bowhead and humpback vocalizations (same features as used in Table II). Correct classification is represented by a colored histogram bin plotted on the background of the same color (e.g., dark blue on light blue). Histograms of all training and validation results are plotted in the same figure. The classifier was able to discriminate between the signal types with 100% accuracy and an $AUC = 1.00$ for all propagation ranges. From these results it is clear that the class distributions for both the bowhead and humpback synthetic signals are not significantly changed by propagation effects (as a function of range). It should be kept in mind that these results provide a *preliminary* look at how propagation impacts the five features incorporated in the

discriminant analysis. At this point the model only produces simplified propagation conditions. In order to make more general statements about the results, additional work needs to be done to ensure that the pulse propagation model is sufficiently representative of true propagation conditions.

Table III: Geoacoustic parameters used to model the sediment half-space. These parameters were selected to be consistent with the silty-clay sediment observed during the Gulf of Mexico experiment.

Parameter	Value
Bottom depth, H	72 m
Density, ρ	1.77 g/cm ³
Compressional sound speed, c_p	1 646 m/s
Shear sound speed, c_s	400 m/s
Compressional attenuation, α_p	0.8 dB/ λ
Shear attenuation, α_s	2.5 dB/ λ

IMPACT/APPLICATIONS

Detection and classification of cetaceans has become critically important to the US Navy due to an ever increasing requirement for environmental stewardship. Passive acoustics continues to be the best method to carry out this task but current techniques provide only a partial solution; most detectors are either too general, leading to unacceptably high false alarm rates, or are too specialized (i.e., species- or location-specific) leading to many missed detections. Furthermore, future military platforms will have to support smaller complements and deal with ever-increasing data throughput, so that automation of on-board systems is essential. In addition, the technique is well suited to autonomous systems since a much smaller bandwidth is needed to transmit a classification result than to transmit raw acoustic data. The success of the machine classifier in discriminating cetacean vocalizations suggests that it could be applied to other passive acoustic classification problems which currently employ human audition. This would be particularly useful if expert listeners aren't available – such as diagnosing heart murmurs in remote communities that lack a cardiologist, or as part of the triage process in a hospital emergency department. Alternatively, the machine classifier is ideally suited when the sheer volume of data makes human audition untenable – such as classifying ocean acoustic data for species population monitoring. Finally, developing a robust classifier for passive marine mammal vocalizations is also a first step to testing the algorithm on passive transients generated by submarines to examine its potential for passive detection and classification of submarines.

RELATED PROJECTS

This research will benefit from DRDC Atlantic's Force ASW Program in which DRDC's aural classification algorithms (including the marine mammal classification algorithm) is being integrated into DRDC's System Test Bed (STB). The STB is used to evaluate sonar algorithms in a military context. Some of the insights to be gained will be: whether the aural classifier can reduce false alarms from marine mammals; does the classifier reduce operator workload required by environmental considerations (the so-called green navy) to enable greater concentration on potential targets; is the aural classifier easily integrated into a navy platform?

REFERENCES

- [1] Robert J. Urick, Principles of Underwater Sound, 3rd ed., McGraw-Hill Book Company, 1983.
- [2] F. B. Jensen, W. A. Kuperman, M.B. Porter, and H. Schmidt, Computational Ocean Acoustics, AIP Series in Modern Acoustics and Signal Processing, 1st ed., American Institute of Physics, (1994).
- [3] Stan Dosso, Peter Giles, Gary Brooke, Diane McCammon, Sean Pecknold, and Paul Hines, “Linear and Nonlinear Measures of Ocean Acoustic Environmental Sensitivity”, *J. Acoust. Soc. Am.* 121 (1), pp. 42-45, (2007).
- [4] Carolyn M. Binder and Paul C. Hines, Automated Aural Classification Used for Inter-species Discrimination of Cetaceans, *J. Acoust. Soc. Am.*, 135 (4), pp. 2113–2125, (2014).
- [5] Victor W. Young and Paul C. Hines, "Perception-based automatic classification of impulsive-source active sonar echoes," *J. Acoust. Soc. Am.*, 122 (3) pp. 1502-1517, (2007).
- [6] Carolyn M. Binder and Paul C. Hines, “Modelling the impact of the ocean environment on automatic aural classification,” *Proceedings: Underwater Acoustics Conference 2014*, Rhodes, GR, (2014).
- [7] Stefan Murphy and Paul C. Hines, “SNR dependence and temporal robustness of an automatic aural classifier,” *Proceedings: 4th International Conference on Underwater Acoustic Measurements*, Kos, Greece, (2011).
- [8] Stefan Murphy and Paul C. Hines, “Examining the robustness of automated aural classification of active sonar echoes”, *J. Acoust. Soc. Am.* 135 (2), pp. 626-636, (2014).
- [9] Paul Etter, Underwater Acoustic Modeling: Principles, Techniques, and Applications, 1st ed., Elsevier Applied Science, (1991).
- [10] Michael J. Buckingham, Ocean-acoustic propagation models, *J. Acoustique*, 43 (30), pp. 233-287, (1992).
- [11] D. K. Mellinger, K. M. Stafford, S. E. Moore, R. P. Dziak, and H. Matsumoto, An overview of fixed passive acoustic observation methods for cetaceans. *Oceanography*, 20: pp. 36–45, (2007).
- [12] Henrik Schmidt, Massachusetts Institute of Technology, Cambridge, Massachusetts, 1999. OASES 2.2. Available: <http://acoustics.mit.edu/faculty/henrik/oases/oases.html>.
- [13] D.M.F. Chapman, “You can’t get there from here: Shallow water sound propagation and whale localization”, *Canadian Acoustics*, 32, pp. 167-171, (2004).
- [14] X. Mouy, M Bahoura, and Y. Simard, “Automatic recognition of fin and blue whale calls for real-time monitoring in the St. Lawrence,” *J. Acoust. Soc. Am.*, 126, pp. 2918-2928, (2009).

PUBLICATIONS

1. Carolyn M. Binder and Paul C. Hines, “Automated Aural Classification Used for Inter-species Discrimination of Cetaceans,” *J. Acoust. Soc. Am.*, 135 (4), pp. 2113–2125, 2014. [published, refereed].

2. Carolyn M. Binder and Paul C. Hines, “Modelling the impact of ocean environment on automatic aural classification of marine mammals,” Proceedings: 2nd International Conference on Underwater Acoustics, Rhodes Greece, June 22-27, 2014. [published].
3. Carolyn M. Binder and Paul C. Hines, “Examining the Impact of the Ocean Environment on Cetacean Classification Using a Propagation Model,” 167th meeting of the Acoustical Society of America, May 2014. [Abstract published].

HONORS/AWARDS/PRIZES

Carolyn Binder (Dalhousie University/Defence R&D Canada – Atlantic):

1. NSERC Postgraduate Scholarship, sponsored by National Science and Engineering Research Council (NSERC);
2. Killam Predoctoral Scholarship, sponsored by Dalhousie University;
3. Young Scientist Award, Awarded for best presentation of a paper by a graduate student, sponsored by the Underwater Acoustics 2014 Conference, Greece.

Table I: Summary of the main considerations when selecting a numerical technique for ocean acoustic pulse propagation modeling. The division between low and high frequencies is set at 500 Hz, following Etter [9]; he selected this threshold somewhat arbitrarily but it does reflect that above 500 Hz some techniques become computationally intensive, and below it the physics of ray-tracing may be ‘questionable.’ The shallow water regime includes all water depths for which the sound significantly interacts with the bottom boundary. The column titled ‘Pulse propagation’ considers numerical techniques that are known to have already implemented a pulse propagation capability. The column labeled ‘non-proprietary’ considers whether there is an available model implementation that is well-established, as well as freely-available. Only a single example implementation is given in this table, although many other implementations exist. The examples listed here are well-known in the ocean acoustics research community. From this table it is evident that each numerical technique has many benefits – the primary motivators for selecting wavenumber integration were the frequency range and the availability of a non-proprietary model with pulse propagation implemented. Note that no explicitly time-domain models were considered in the review since according to [2], most problems of practical importance in underwater acoustics (long-range propagation of pulses of ‘finite’ bandwidth) favor the Fourier synthesis technique.

Parameter	Frequency Range	Suitable for shallow water	Suitable spatial scale	Range-dependent	Pulse propagation	Non-proprietary	Example
<i>Ray-tracing</i>	high	yes*	yes	fully	yes	yes	BELLHOP
<i>Normal mode</i>	low to few kHz	yes	yes	with approximations [†]	unknown	yes	KRAKEN
<i>PE</i>	low	yes	yes	fully	unknown	yes	RAM
<i>Wavenumber integration</i>	low to few kHz	yes	yes	with approximations [†]	yes	yes	OASES

* Given a sufficiently high frequency signal.

[†] I.e., range-dependence can be implemented using a suitable approximation – either coupled-modes or adiabatic approximation.