MODELLING THE IMPACT OF OCEAN ENVIRONMENT ON AUTOMATIC AURAL CLASSIFICATION OF MARINE MAMMALS

Carolyn M. Binder[a,b], Paul C. Hines[a]

[a]Department of Oceanography, Dalhousie University, LSC Ocean Wing, 1355 Oxford St., PO Box 15000, Halifax, NS, Canada, B3H 4R2
[b]Defence Research and Development Canada, PO Box 1012, Dartmouth, NS, Canada, B2Y 3Z7

Carolyn Binder, DRDC – Atlantic Research Centre, PO Box 1012, Dartmouth, NS, Canada, B2Y 3Z7. Fax: 1 (902) 426-9654. E-mail: carolyn.binder@drdc-rddc.gc.ca

Abstract: *Passive acoustic monitoring (PAM) is widely in use to study marine mammals in their underwater habitats. Since marine mammals can be found in all ocean basins, their habitats cover diverse underwater environments. Properties of the ocean environment, such as sound speed profile, bathymetry, and sediment properties can be markedly different between these diverse habitats, leading to differences in how a marine mammal's vocalization is altered by propagation effects. This distortion of vocalizations may impact the accuracy of PAM systems. Thus, to develop a PAM system capable of operating in numerous environments one must understand how propagation effects impact these systems.*

Previous effort has shown that a prototype aural classifier developed at Defence R&D Canada could successfully discriminate several cetacean species' vocalizations in a relatively limited data set. The aural classifier was found to be an effective PAM tool because it employs perceptual signal features, which model features used by the human auditory system. The current work used the OASES (Ocean Acoustics and Seismic Exploration Synthesis) pulse propagation model to examine the robustness of the classifier under various environmental conditions. Preliminary results from transmitting cetacean vocalizations over several ranges in a simulated underwater environment are discussed. The modelled environment used to obtain these results was based on environmental data collected during propagation trials. Aural classification accuracy was compared for signals propagated over different ranges and provided a preliminary measure for the robustness of the perceptual features to propagation effects.

Keywords: *Aural classifier, marine mammals, passive acoustic monitoring, propagation*

DRDC-RDDC-2014-N32

1. INTRODUCTION

Properties of the ocean environment – such as sound speed profile (SSP), bathymetry, sediment properties, and ambient noise profiles – can be markedly different between regions where passive acoustic monitoring (PAM) is used to observe cetaceans. This leads to differences in sound propagation characteristics and results in distortion of received cetacean vocalizations [1]. The accuracy of PAM systems may be affected if this is not accounted for; however, little research has yet been directed towards this problem. Helble *et al*. [2] demonstrate the significant impact of the ocean environment on PAM. By using a parabolic equation propagation model to simulate calls originating from sources at various locations with respect to a fixed receiver, they were able to show that the probability of detecting a humpback whale call is environment-dependent and can be markedly different between monitoring locations. Based on these results, it may be concluded that a thorough understanding of how the environment impacts the signal features used for detection and/or classification is required to develop an automatic recognition system capable of operating effectively under numerous environmental conditions.

Previous results show that a prototype computer-based aural classifier developed at Defence Research and Development Canada (DRDC) [3] can be used to successfully discriminate vocalizations from several cetacean species [4]. The success of the aural classifier is due to the perceptual signal features it employs. These are different than features obtained using conventional signal processing techniques because they take into account how a listener perceives sound [3, 4].

The current work aims to determine the robustness of the aural classifier under various environmental conditions through a combination of empirical measurements and simulated results. This paper focuses on results obtained by using the Ocean Acoustics and Seismic Exploration Synthesis (OASES) propagation model [5] to simulate synthetic bowhead and humpback vocalizations propagated over ranges of 0 to 20 km. The properties used to characterize the modelled environment were based on environmental data collected during a propagation experiment that took place in the Gulf of Mexico during the spring of 2013.

2. DATASET

Recordings of example bowhead and humpback vocalizations were obtained from the MobySound website [6]; however, these calls were subject to unknown propagation effects when they were recorded. Therefore, synthetic signals were developed to provide known starting signals with no propagation effects applied prior to the experiments. The synthetic signals were based on an example set of high signal-to-noise ratio recordings of bowhead song endnotes and humpback song units. To generate the synthetic signals, the noise was reduced in each example call using wavelet analysis. Then the mean signal and empirical orthogonal functions (EOFs) were calculated from the wavelet-transformed vocalizations for each species [7]. To generate a single synthetic signal, random weights were applied to the EOFs that contained 95% of the variance, the weighted EOFs were added to the mean signal, and finally the inverse wavelet transform was performed. Applying different randomized weights to the EOFs for each synthetic signal generated a set of 155 signals per species. The resulting collection of synthetic signals had similar mean and variance for the perceptual features that were considered most important for

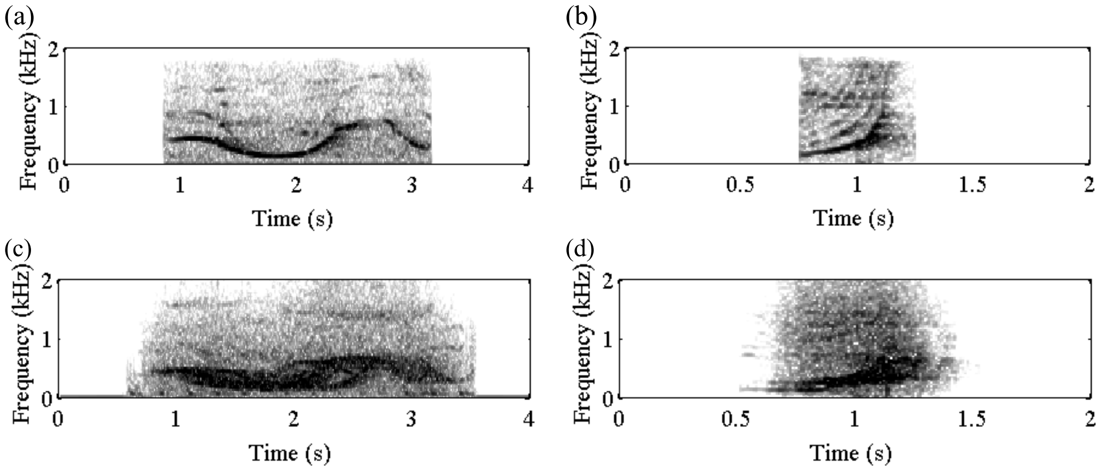

Fig.1: Example spectrograms of (a) bowhead, (b) humpback, (c) synthetic bowhead, and (d) synthetic humpback vocalizations.

discriminating between the example bowhead and humpback vocalizations. Fig. 1 shows example spectrograms of the real and synthetic bowhead and humpback vocalizations.

3. METHODS

A two-day sea trial was conducted in the Gulf of Mexico, approximately 74 km south of Panama City, FL, from 30 April to 1 May 2013. A pictorial summary of the experimental setup is given in Fig. 2 (a). Two moorings were deployed, each with two hydrophones at different depths within the water column. Then both real and synthetic vocalizations were transmitted to the moored receivers from a source deployed from CFAV QUEST while the ship drifted. After the transmissions were completed (approximately one hour), the range from the moored receivers was increased and the transmissions were repeated. Environmental properties were measured throughout the experiments in order to understand the propagation conditions, and to provide realistic parameters for modelling. CTD casts were performed at each location the signals were transmitted to characterize water column properties. Information on the sediment characteristics was obtained from several Free Fall Cone Penetrometer (FFCPt) casts along the ship's track [8]. Data from the FFCPt provided information on the sediment type (e.g., silty-clay, sand), from which the relevant geo-acoustic parameters were estimated.

To simulate this experimental procedure, the pulse propagation module of OASES [5], referred to as OASES-OASP, was used to propagate signals through a range-independent environment. OASES is a general-purpose computer code that uses the wavenumber integration method for modelling seismoacoustic propagation in horizontally stratified waveguides. OASES-OASP calculates the depth-dependent Green's function and determines the acoustic transfer function at each receiver by evaluating the wavenumber integral [5]. The model outputs the received pressure time series for each source signal.

The signals propagated with OASES-OASP were input to the aural classifier algorithm (note that the signals recorded during the experiments are not considered in this paper). The aural classification process is broken into three phases: First, a simple auditory model is applied. Second, the perceptual signal features are calculated for each signal. Finally, a Bayesian classifier is applied [3, 4]. The classifier was *trained* with signals propagated

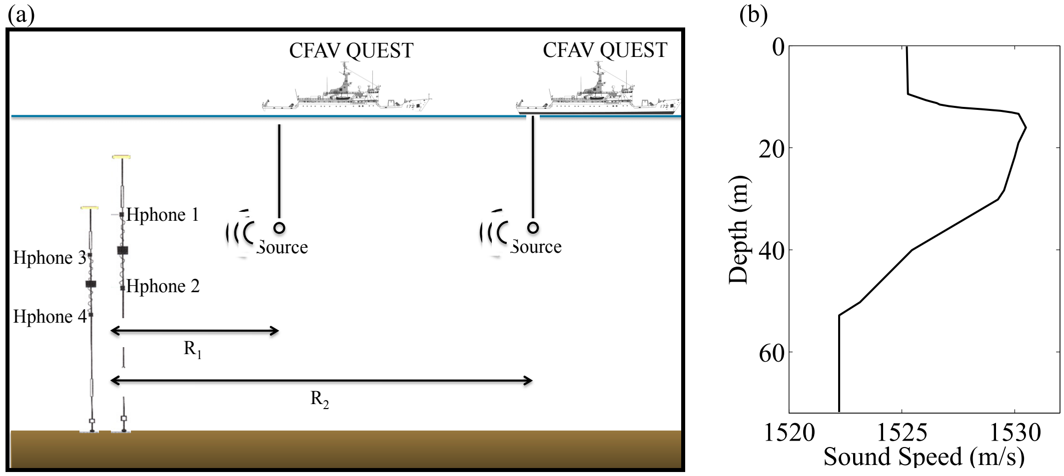

Fig.2: (a) Representation of the experimental setup. The ship first deployed two hydrophone moorings, moved to the first location and transmitted the set of signals, then moved further away from the recorders and retransmitted the signals. R_1 and R_2 represent the horizontal range the signals propagated from source to the midpoint between the moorings. (b) Sound speed profile measured during the morning of 30 April 2013. This was used to characterize the water column for modelling with OASES-OASP.

over 0 km in the model (i.e., a receiver was placed coincident with the source to capture any processing effects in the model), for which the classifier was provided the class labels. This classifier was then *tested* with signals propagated in the model over 5, 10 and 20 km ranges, for which the classifier had no direct knowledge of the associated class labels.

Classification accuracy and area under the ROC curve, *AUC*, were used to evaluate classifier performance. The *AUC* can vary between a value of 1.00, indicative of an ideal classifier, and 0.50, equivalent to randomly assigning a classification decision. In this way, OASES-OASP was used in conjunction with the aural classifier to study the effects of propagation on the perceptual features as a function of range between source and receiver.

4. RESULTS AND DISCUSSION

The model was run for geometries and environmental parameters consistent with the Gulf of Mexico sea trial. The environment was modelled as a water column overlaying a sediment half-space. The geo-acoustic parameters of the sediment half-space (density = 1.77 g/cm^3, compressional sound speed = 1646 m/s, shear sound speed = 400 m/s, compressional attenuation = 0.8 dB/λ, and shear attenuation = 2.5 dB/λ) were consistent with the silty-clay sediment type observed at the experimental site. The sound speed profile shown in Fig. 2 (b) was measured the morning of 30 April 2013 and was used to model the water column characteristics. The following results considered only propagation of the synthetic signals (real vocalizations will be analyzed in future work) with the source and receiver located at depths of 30 m and 36 m, respectively.

Fig. 3(a) shows results obtained by training the classifier on synthetic signals propagated with OASES-OASP over a range of 0 km. The remaining panels in Fig. 3 depict the results of testing the classifier on signals propagated with the model over ranges of 5 to 20 km. These plots show histograms of projected feature values that were obtained from a linear combination of the five features (duration, global mean subband decay time,

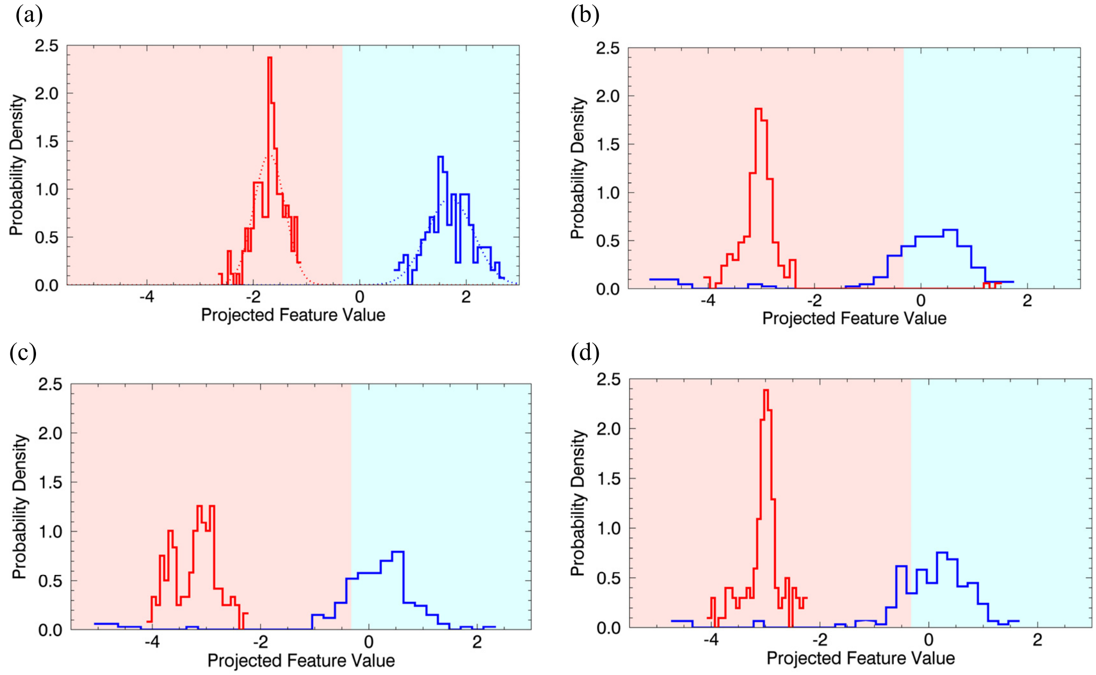

Fig.3: Classification results for synthetic bowhead (blue) and humpback (red) signals. (a) The classifier was trained using signals that were propagated using OASES-OASP over 0 km range. The dashed lines in panel (a) represent the theoretical Gaussian PDFs used to determine the decision regions. The classifier was tested on signals that were propagated over ranges of (b) 5 km, (c) 10 km, and (d) 20 km.

Transmission Range [km]	Accuracy	*AUC*	Mean (Bowhead)	Variance (Bowhead)	Mean (Humpback)	Variance (Humpback)
0	100%	1.00	1.70	0.19	-1.70	0.09
5	91%	0.97	-0.13	1.98	-3.01	0.33
10	92%	0.95	-0.01	1.28	-3.22	0.17
20	88%	0.97	-0.03	1.01	-3.04	0.12

Table 1: Accuracy, AUC, and class means and variances for the synthetic vocalizations.

local maximum subband decay time, frequency of global maximum subband attack time, and peak loudness value [3, 4]) that best discriminated between the real vocalizations from MobySound. Correct classification is represented by a coloured histogram bin plotted on the background of the corresponding colour (e.g., dark blue on light blue). Quantitative results are listed in Table 1.

The classifier was able to discriminate between the signals in the training set with 100% accuracy and *AUC* of 1.00. There was a trend for accuracy to decrease as the propagation range increased, although the *AUC* values showed little change. The *AUC* values were similar because the separation between the bowhead and humpback classes was maintained. These two trends are reflected in the decision regions in panels (b) through (e) of Fig. 3; in each there was little overlap between class distributions, though the distributions shifted with respect to the decision threshold (i.e., the boundary separating the red and blue backgrounds). The variance of both the bowhead and humpback classes changed with propagation range. These preliminary results suggested that propagation conditions affect at least one of the five perceptual signal features used for classification.

5. CONCLUSIONS

The OASES-OASP propagation model was used to propagate synthetic signals through a modelled environment based on environmental properties measured during a sea trial performed in the Gulf of Mexico. Preliminary classification results indicate that the environment may impact some of the perceptual signal features. Further work is required to quantify how robust each of the perceptual features is to propagation effects and which environmental properties most affect the features. Future work also includes comparing results from propagation models to results obtained from the sea trials.

6. ACKNOWLEDGEMENTS

The authors would like to thank Dr. Sean Pecknold and Mr. Stefan Murphy of DRDC, and Dr. Tetjana Ross of Dalhousie University for advice and insights. The authors wish to acknowledge MobySound for the availability of marine mammal vocalization data. This work was supported in part by the U.S. Office of Naval Research, Code 32, Marine Mammals & Biological Oceanography Program under Grant N00014-14-1-0237.

REFERENCES

[1] **F. B. Jensen, W. A. Kuperman, M. B. Porter, and H. Schmidt**, *Computational Ocean Acoustics*, R. Beyer, American Institute of Physics, pp. 3–10 & 203–270, 1994.

[2] **T. A. Helble, G. L. D'Spain, J. A. Hildebrand, G. S. Campbell, R. L. Campbell, and K. D. Heaney**, Site Specific Probability of Passive Acoustic Detection of Humpback Whale Calls From Single Fixed Hydrophones, *J. Acoust. Soc. Am*, 134 (3), pp. 2556–2570, 2013.

[3] **V. W. Young and P. C. Hines**, Perception-based Automatic Classification of Impulsive-Source Active Sonar Echoes, *J. Acoust. Soc. Am*, 122 (3), pp. 1502-1517, 2007.

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

[5] **H. Schmidt**, OASES Version 3.1 User Guide and Reference Manual, *Ocean Acoustics Library* (2011), http://oalib.hlsresearch.com/FFP/index.html (Viewed 7 Apr. 2014).

[6] **S. Heimlich, D. Mellinger, and H. Klinck**, *The MobySound Database for Research in Automatic Recognition of Marine Mammal Calls* (2012), http://www.mobysound.org (Viewed 24 Mar. 2014).

[7] **L. R. LeBlanc and F. H. Middleton**, An Underwater Acoustic Sound Velocity Data Model, *J. Acoust. Soc. Am.*, 67 (6), pp. 2055–2062, 1980.

[8] **J. Osler, A. Furlong, and H. Christian**, A Sediment Probe for the Rapid Assessment of Seabed Characteristics, In *Acoustic sensor techniques for the shallow water environment*, Ischia, Italy, A. Caiti, N. R. Chapman, J. P. Hermand, and S. M. Jesus, pp. 171–181, 2006.