A DESCRIPTION OF THE WORKSHOP DATASETS

Francine Desharnais[1], Marjo H. Laurinolli[2], Douglas J. Schillinger[3], Alex E. Hay[3]
[1]DRDC Atlantic, PO Box 1012, NS, Canada, B2Y 3Z7, francine.desharnais@drdc-rddc.gc.ca
[2]Department of Fisheries and Oceans, Bedford Institute of Oceanography, PO Box 1006, Dartmouth, NS B2Y 4A2
[3]Department of Oceanography, Dalhousie University, Halifax, NS, B3H 4J1

ABSTRACT

An acoustic dataset was provided to the participants of the *2003 Workshop on detection and localization of marine mammals using passive acoustics* (Dartmouth, NS, Canada, 19-21 November 2003). This document contains environmental and other technical information regarding supporting the acoustic dataset.

RÉSUMÉ

Un ensemble de données a été mis à la disposition des participants de l'*Atelier 2003 sur la détection et la localisation des mammifères marins à l'aide du repérage acoustique passif* (Dartmouth, NÉ, Canada, 19-21 novembre 2003). Ce document contient de l'information sur l'environnement, et autres détails techniques reliés aux données acoustiques.

1. INTRODUCTION

In support of the *2003 Workshop on detection and localization of marine mammals using passive acoustics* (Dartmouth, NS, Canada, 19-21 November 2003), an initial dataset was provided for the participants to test their algorithms. This DRDC/Dalhousie dataset was originally limited to North Atlantic right whale sounds recorded in the Bay of Fundy during 2000 and 2002. After initial distribution of this dataset, an additional dataset was made available by the Cornell Laboratory of Ornithology. This second dataset was made of similar recordings of right whales off the US east coast.

This document contains a brief description of the datasets as given to the participants. It also contains environmental information for the Bay of Fundy to support the data analysis efforts. The complete DRDC/Dalhousie package can be downloaded from the workshop web site [1]. The Cornell dataset is available on request (see Sec. 4).

2. DRDC/DALHOUSIE DATASET

2.1 The Location

This dataset was collected in the Bay of Fundy, Canada, in September 2002, using 5 Ocean Bottom Hydrophones (OBHs). Figure 1 shows the OBH locations, as well as water depth contours in 50-m increments. The OBH pattern is approximately 14.5 km wide – this distance was selected based on previous localization experiments in the Bay.

2.2 The Recording Devices

The OBHs (Fig. 2) are self-recording devices that are moored to the seabed. Each OBH system includes an omnidirectional hydrophone (OAS model E-2SD), a canister with electronics to amplify and record the hydrophone signal, and an acoustic release. The units recorded files of 19.58 minutes in length, with a variable spacing of approximately 10 seconds between files during which no data were recorded. The sampling frequency is programmed before each deployment, and affects the overall recording duration. The hydrophone is 0.9 m from the seabed when deployed.

Table 1 includes the location of the ship and water depth where the OBHs were deployed. The OBHs were deployed on 2 separate days, during different phases of the tide. Because the tidal currents in this area are strong, up to 1 m/s during the deployment, and the OBH descent rate is of similar magnitude, the actual position of the OBH on the bottom could differ from the ship's position by up to 100 m.

Table 1. Location of the ship during the deployment of the five OBHs in September 2002.

OBH	Deployment position		Water depth (m)
	Latitude (N)	Longitude (W)	
C	44.60202	66.49737	210
E	44.60197	66.31650	134
L	44.66320	66.40397	183
H	44.72943	66.31638	123
J	44.72910	66.49693	170

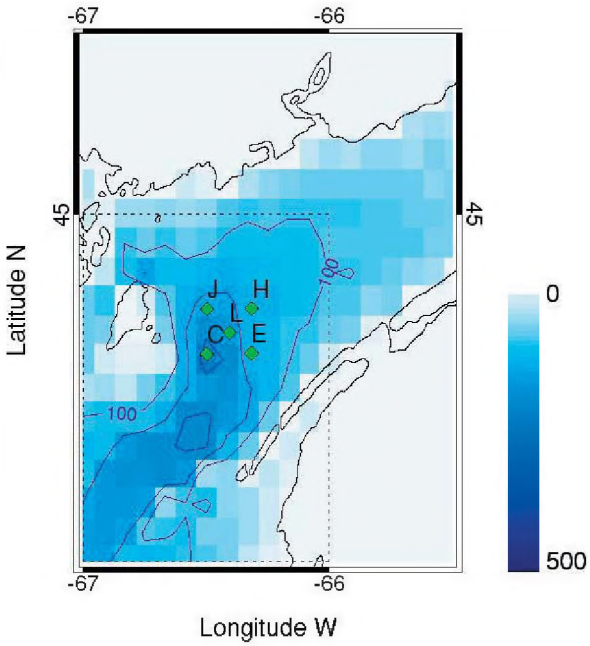

Figure 1. Locations of the five OBHs in the Bay of Fundy.

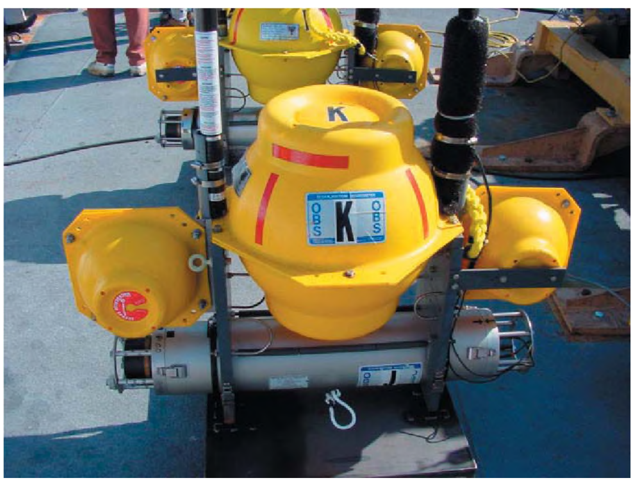

Figure 2. An OBH ready for deployment from the research ship.

To better refine their positions, a hydrophone was deployed from the ship at several locations around each OBH, while the ship was drifting. The output from this hydrophone was synchronized to the GPS clock. Each OBH is equipped with a pinger that sends out a series of acoustic signals at known times relative to the OBH clock, which is also synchronized to the GPS time base prior to deployment. The arrival times of these signals at the hydrophone give the travel time between the OBH and the hydrophone. Travel times recorded at each OBH (order of 40 per OBH) were reduced to 4 mean values at 4 mean locations. The exact OBH location is determined from the intersections of the four corresponding circles. The results for OBH C are shown in Figure 3. The OBH localization experiment led to the updated positions listed in Table 2, along with estimates of position uncertainty (the standard deviation of the intersections among the equal travel time circles, Fig. 3).

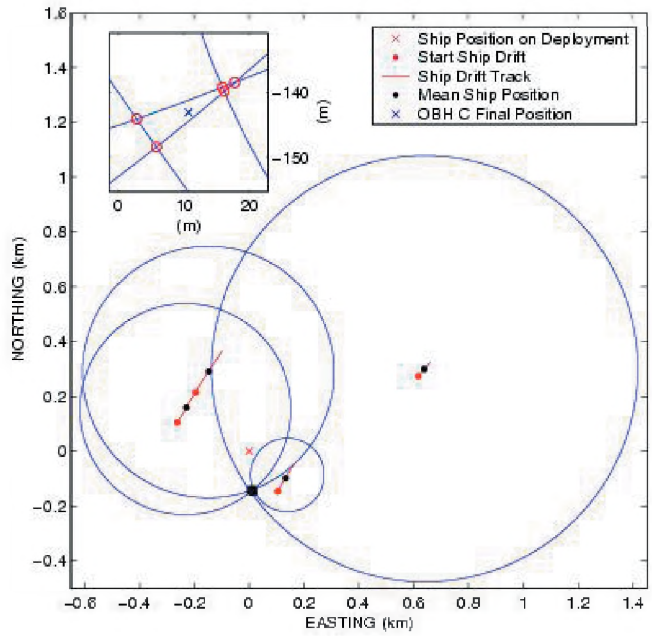

Figure 3. Results of localization experiment for OBH C. Red lines indicate the ship drift tracks, solid red circles the drift start, and solid black circles the mean ship position during the drift. The red 'x' indicates the ship's position when the OBH was launched. The solid black square is expanded in the inset and shows the intersections (open red circles) among the equal travel time circles (in blue) about the mean ship positions, and the best estimate position of the OBH on the seabed (blue x). Note the 150 m distance (approximate) between this final position and the position at launch.

Table 2. OBH updated positions and uncertainty.

OBH	Updated positions		Uncertainty	
	Latitude (°N)	Longitude (°W)	Northing (m)	Easting (m)
C	44.60073	66.49723	2.15	6.06
E	44.60237	66.31591	5.13	4.16
L	44.66203	66.40453	3.14	2.08
H	44.73051	66.31556	12.55	11.47
J	44.73038	66.49619	0.42	9.72

Each OBH has an accurate temperature-stabilized clock. Clock drift was estimated by measuring the offset relative to GPS time before the deployment and after the recovery. Table 3 lists the clock offset for each OBH deployed in September 2002, as measured before and after the deployment. A negative number means the OBH lags the GPS time signal. Linear drift is assumed over the deployment duration. The acoustic files provided within this data set have not been corrected for clock drift.

Table 3. Clock offsets for each OBH in the 2002 experiment.

OBH	Synchronization time	Clock offset [ms]
	Before deployment	
C	9 Sep 20:16	1.888
E	9 Sep 16:53	0.0366
L	10 Sep 15:55	-2.042
H	9 Sep 21:14	4.139
J	9 Sep 18:46	-1.985
	After recovery	
C	18 Sep 11:48	-3.142
E	18 Sep 12:57	1.809
L	18 Sep 16:24	0.337
H	18 Sep 14:02	-9.437
J	18 Sep 11:34	34.262

To maximize the OBHs bottom time, a sampling frequency of 1200 Hz was selected in 2002, with a low-pass filter of 800 Hz. A filter roll-off frequency above the Nyquist frequency was selected to maximize localization opportunities for sounds in the upper end of the frequency range. Calibration files are included with the data set, but users should keep in mind the energy fold over. The amplifiers in all electronics packages were bench-calibrated, while the hydrophones were calibrated at the DRDC Atlantic calibration barge. The OBHs have a constant sensitivity over 50 to 700 Hz, and the sensitivity is independent of the instrument or the hydrophone, except for OBH-J, which recorded levels approximately 20 dB lower than expected. The data were digitized using a 12-bit A/D converter with ±5 V range. The levels can be calibrated using Equation (1):

$$L = P + 20 \cdot LOG_{10}[5/(2^{12}-1)] - G - S \quad (1)$$

where L is the calibrated level in dB re 1 μPa^2, P is the initial level in dB re 1 V^2, G is the gain [59 dB], and S is the hydrophone sensitivity [-187 dB re 1 V/μPa].

2.3 The Environment

The bathymetry for the environment was extracted from the DBDB-V database (version 4.2) from the Naval Oceanographic Office, accessed through their web site [Idbms.navo.navy.mil/dbdbv/dbvquery.html]. The data were approved for release, and have a 5-min resolution in both latitude and longitude. The data set includes the file *dbdbv.txt*, which contains the depth in m over the 5-min grid from W to E, N to S. The file *dbdbv_coord.txt* contains the same information with the coordinates for each point.

Several expendable bathythermograph (XBTs) and conductivity-temperature-depth (CTD) profiles were taken during the experiment. Table 4 lists the profiles nearest in time to each data sample. Figure 4 shows the selected collection of temperature profiles. These data are included as ASCII files in the data set.

Table 4. XBT and CTD records close in time to the data samples included in the data set.

Data Type	Filename	Date [2002]	Time (UTC)	Lat (°N)	Long (°W)	Water Depth (m)
XBT	T7_00004	Sept 11	19:28	44 40.379	66 26.967	194
XBT	T7_00005	Sept 11	23:34	44 38.601	66 29.634	202
XBT	T7_00007	Sept 12	11:41	44 38.611	66 31.354	195
XBT	T7_00008	Sept 12	18:42	44 40.566	66 27.038	192
CTD	q269ctd3	Sept 11	19:48	44 40.411	66 27.263	194
CTD	q269ctd4	Sept 12	12:33	44 36.020	66 29.875	215.1
CTD	q269ctd5	Sept 12	18:41	44 40.479	66 27.384	192.7
CTD	q269ctd6	Sept 13	11:53	44 39.770	66 29.020	197.5
CTD	q269ctd7	Sept 13	12:14	44 40.000	66 28.063	193.8
CTD	q269ct18	Sept 14	9:17	44 40.484	66 26.115	190.1

The sediment in the area is composed mainly of a layer of variable thickness LaHave clay over a thick layer of Scotian Shelf drift [2]. The postglacial silty sandy clay is loosely compacted, and is generally characterized by a low compressional sound speed. The physical properties of the LaHave clay found on the Scotian Shelf around Nova Scotia are typically those listed in Table 5 [3]. The Scotian Shelf drift is glacial till, a cohesive, poorly sorted sediment, generally containing angular fragments in the pebble/cobble/boulder range. It is predominantly sandy, but contains abundant silt and clay [3]. Its typical properties are also listed in Table 5; the compressional sound speed, attenuation and density were taken from [3], shear speed and attenuation were estimated from [4].

The thickness of the upper sediment layer was extracted from sub-bottom profiler data for the area. The file *sediment_thickness.dat* contains our estimates of sediment thickness as a function of latitude and longitude along specific tracks. Note that the thickness data in this file are for the upper clay sediment layer only. Figure 5 shows the sediment thickness data. Note that the layer thickness is expected to increase north of the OBHs. A poorly-defined reflector led to estimates at end "3" that are lower than expected. It is believed that gas within the surficial sediment layer led to the lower (and possibly inaccurate) estimates over this small area, though the data were insufficient to confirm this.

Table 5. Seabed parameters for upper sediments.

	LaHave clay	Scotian Shelf drift
Compressional sound speed (km/s)	1.261-1.49	1.745-1.92
Comp. attenuation (dB/m-kHz)	0.023, 0.056	0.0065
Density (g/cm^3)	1.5-1.54	2.1
Shear speed (km/s)	0.0	0.4-0.5
Shear attenuation (dB/m-kHz)	0.0	10.0

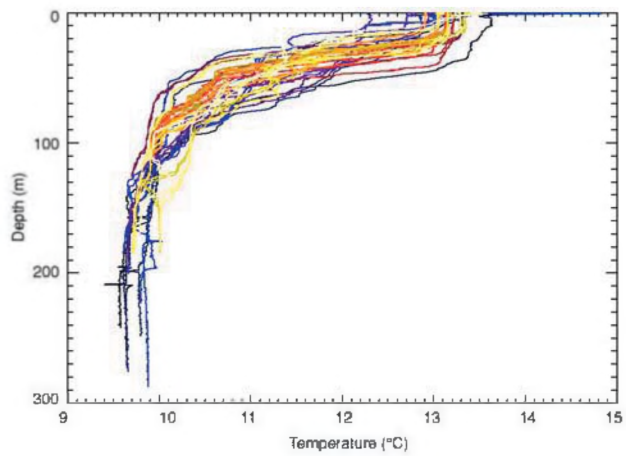

Figure 4. Temperature profiles from XBT and CTD records.

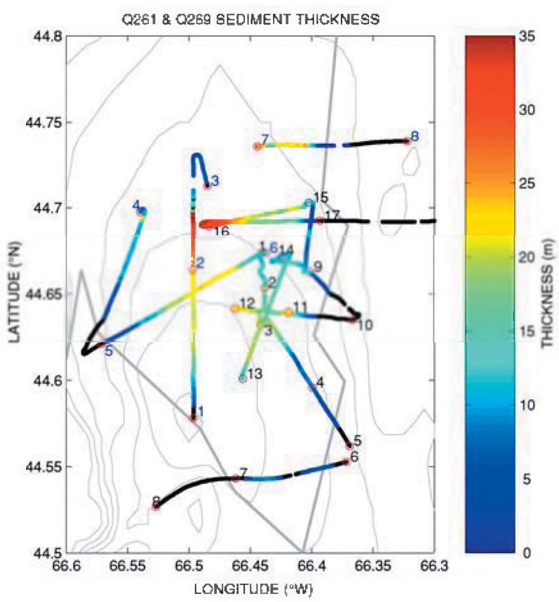

Figure 5. Upper sediment layer thickness.

2.4 The Acoustic Data

The raw OBH data were converted to audio files of WAV format using the READ_WAV function of IDL (Interactive Data Language, Research Systems Inc). The READ_WAV function does not normalize the data, which are integers within the file. Various sound types – vocalizations and "gunshot" sounds – were selected, and isolated into five 30-sec files (one for each OBH). The 30-sec period was selected so that the sound recorded on OBS-L was approximately in the middle of the file. For a given sound, all five files start at precisely the same time, except for the clock drift which has not been accounted for. Table 6 includes a list of all filenames, with a brief description of the sound types and the file start times. The filenames in Table 6 are to be completed with the letter describing each OBH; for example, filename 'S013-1" indicates that there are 5 files named: 'X-S013-1.wav', where 'X' will be either C, E, L, H or J, depending on which OBH the file was recorded on.

For the purpose of testing detection algorithms, an 18-min long file (L-138.wav) from OBS-L is included. This file was selected as it was rich in sound occurrences and sound types. An equivalent file for the other OBHs can be made available on request.

Table 6. List of sounds in the data set.

	Filename	File start date [2002]	File start time	Sound type
1	S013-1	11 Sep	17:23:04	Gunshot
2	S035-2	12 Sep	0:34:23	Gunshot
3	S070-3	12 Sep	12:00:44	Gunshot
4	S093-4	12 Sep	19:24:09	Gunshot
5	S110-5	13 Sep	1:03:06	Gunshot
6	S092-7	12 Sep	19:08:57	Low-frequency call
7	S093-9	12 Sep	19:32:15	Low-frequency call
8	S131-10	13 Sep	7:47:15	Low-frequency call
9	S131-11	13 Sep	7:47:37	Low-frequency call
10	S131-12	13 Sep	8:02:16	Low-frequency call
11	S131-13	13 Sep	8:03:04	Low-frequency call
12	S134-6	13 Sep	8:50:56	Low-frequency call
13	S143-8	13 Sep	11:52:50	Low-frequency call
14	S209-14	14 Sep	9:29:52	Mid-frequency call
15	S210-15	14 Sep	9:34:30	Mid-frequency call
16	L-138	13 Sep	10:02:32	Multiple calls

2.5 The Calibration Dataset

A calibration dataset was included for the workshop, though it was taken from a different trial than the acoustic dataset described above. This trial occurred during August 2000, and only four OBHs were available for that trial. The positions of the four OBHs are listed in Table 7. No OBH localization experiments were done in 2000. The sampling frequency was 5000 Hz, with a low-pass filter of 1000 Hz.

Table 7. Location of the ship during the deployment of the four OBHs in August 2000.

OBH	Deployment position		Water depth (m)
	Latitude (N)	Longitude (W)	
B	44.7118	66.3494	131
C	44.6714	66.3753	165
D	44.6664	66.4331	190
E	44.7065	66.4083	166

The dataset was believed to be composed of transmissions of right whale calls played back with a projector lowered into the water from a rigid-hull inflatable boat (RHIB). These transmissions were made by Susan Parks (Woods Hole Oceanographic Institute), from a sound file provided by Scott Kraus, of the New England Aquarium. The source was at 20-m depth. Table 8 has the GPS position of the RHIB at two times close to the transmissions. The source level was approximately 155-160 dB re 1 μPa.

Table 8. RHIB position during transmissions.

Time	Latitude (N)	Longitude (W)
17:19:08	44°41.744	66°22.635
17:28:06	44°41.677	66°22.566

The files from 2000 were 70-sec long. The selected files are listed in Table 9, along with the time of selected calls that were suggested for benchmarking purposes. The times listed in Table 9 are the arrival times (relative to the beginning of the file) of the calls on OBH-D. The filenames in Table 9 are to be completed with the letter describing each OBH; for example, filename "S282" indicates that there will be 4 files named: 'X-S282.wav', where 'X' will be either B, C, D or E, depending on which OBH the file was recorded on. The OBH clock drift was negligible in 2000, as the deployment times were short.

Table 9. List of sounds in the calibration data set.

	File name	File start date [2000]	File start time	Relative call time
1	S282	27 Aug	17:20:02	4
2	S282	27 Aug	17:21:44	62
3	S288	27 Aug	17:29:51	63
4	S289	27 Aug	17:31:02	53

It should be noted that the calls selected for the calibration datasets were believed to be from the playback transmissions. This assumption came from the initial localization exercise on these sounds, as the original playback tape was not available when the dataset was created. We cannot rule out that some of the calls from the calibration dataset files may be actual right whale vocalizations, possibly made in response to the playbacks.

The sound speed profile for the 2000 calibration data should be extracted from the XBTs or CTD of Table 10.

Table 10. XBT and CTD records close in time to the 2000 data.

Data Type	File name	Date [2000]	Time (UTC)	Lat (°N)	Long (°W)	Water Depth (m)
XBT	Q253_T 7_00006	Aug 27	14:39	44 42.46	66 20.96	134
XBT	Q253_T 7_00007	Aug 27	18:53	44 41.37	66 20.68	115
CTD	Q253018	Aug 27	06:38	44 40.82	66 24.21	176

3. THE CORNELL DATASET

The Cornell dataset is composed of two sub datasets of approximately 4 hours each: CCB 2001 and GSC 2000. These datasets are copyrighted by the Cornell Laboratory of Ornithology (all rights reserved).

The CCB 2001 data were collected during 2001 in Cape Cod Bay, Massachusetts. The original recordings were made with three retrievable bottom-mounted hydrophones (pop-ups), operating continuously from 8 March, 2001 to 2 April, 2001. Four hours of data were made available for the workshop.

The GSC 2000 data were recorded during 2000 in the Great South Channel (off Cape Cod Bay); these recordings were made for the International Fund for Animal Welfare. Six retrievable bottom-mounted hydrophones were used in 2000. The original recordings were made from 2300 h on 14 May to 2400 h on 14 June, 2000. Again, four hours of data were made available for the workshop.

The Cornell datasets consist of a collection of 5-min long sound files in the AIFF format. The data from all sensors were combined in these files; each sensor is on a separate channel, for a total of 3 channels in 2001, or 6 channels in 2000. The naming convention for the data files is: YYLLDD_ HHMMSS.aif, where YY is the year, LL the month, DD the date, HH the hour, MM the minute and SS the second; this identifies the start time of the data in each file (all times are Greenwich Mean Time). The sampling rate was 2000 Hz for both years.

The surface deployment locations of the hydrophones were based on shipboard GPS readings, as listed in Table 11.

Table 11. Hydrophone positions in 2000 and 2001.

Hydro-phone	2001 Lat [°N]	2001 Long [°W]	2000 Lat [°N]	2000 Long [°W]
1	41.931	70.166	41.844	69.301
2	41.957	70.166	41.858	69.269
3	41.947	70.193	41.830	69.269
4			41.935	69.094
5			41.910	69.095
6			41.923	69.062

4. CONTACTS

The DRDC/Dalhousie dataset is available on the Workshop web site, or by contacting:

Francine Desharnais
DRDC Atlantic
PO Box 1012
Dartmouth,NS,Canada
B2Y 3Z7

The Cornell dataset is available by contacting:

Bioacoustics Research Program
Christopher W. Clark, director
Cornell University Lab of Ornithology
159 Sapsucker Woods Rd.
Ithaca, NY 14850

REFERENCES

[1] The workshop data set can be downloaded from the web site:www.atlantic.drdc-rddc.gc.ca/newsevents/workshop_01_e.shtml

[2] Surficial geology – Eastern Gulf of Maine and Bay of Fundy. Map 4011-G published by the Canadian Hydrographic Service, 1977.

[3] J.C. Osler, “A geo-acoustic and oceanographic description of several shallow water experimental sites on the Scotian Shelf”, DREA Technical Memorandum 94/216, 1994.

[4] E.L. Hamilton, “Acoustic properties of sediments”, Acoustics and Ocean Bottom, A. Lara-Sáenz, C. Ranz-Guerra and C. Carbó-Fité, eds., C.S.I.C., Madrid, pp. 3-58, 1987.

ACKNOWLEDGEMENTS

The authors would like to thank Dr Keith Louden, Dalhousie University, for providing the OBH instruments that were used in the Bay of Fundy, and Dr. Christopher Clark of the Cornell Laboratory of Ornothology for making their dataset available to the workshop participants. The Department of Fisheries and Oceans was also very helpful in making co-author Marjo Laurinolli available to process the DRDC/Dalhousie dataset.