

PERFORMANCE OF SPECTROGRAM CROSS-CORRELATION IN DETECTING RIGHT WHALE CALLS IN LONG-TERM RECORDINGS FROM THE BERING SEA

Lisa M. Munger¹, David K. Mellinger², Sean M. Wiggins¹, Sue E. Moore³, John A. Hildebrand¹

1-Scripps Institution of Oceanography, 9500 Gilman Drive mailcode 0205, La Jolla, CA, 92093-0205, U.S.A.

2-Oregon State University and NOAA Pacific Marine Environmental Laboratory, 2030 SE Marine Science Drive, Newport, OR, 97365, U.S.A.

3-NOAA Alaska Fisheries Science Center and Applied Physics Laboratory, University of Washington, 1013 NE 40th St., Seattle, WA 98195, U.S.A.

ABSTRACT

We investigated the performance of spectrogram cross-correlation for automatically detecting North Pacific right whale (*Eubalaena japonica*) calls in long-term acoustic recordings from the southeastern Bering Sea. Data were sampled by autonomous, bottom-mounted hydrophones deployed in the southeastern Bering Sea from October 2000 through August 2002. A human analyst detected right whale calls within the first month (October 2000) of recorded data by visually examining spectrograms and by listening to recorded data; these manual detections were then compared to results of automated detection trials. Automated detection by spectrogram cross-correlation was implemented using a synthetic kernel based on the most common right whale call type. To optimize automated detection parameters, the analyst performed multiple trials on minutes-long and hour-long recordings and manually adjusted detection parameters between trials. A single set of optimized detection parameters was used to process a week-long recording from October 2000. The automated detector trials resulted in increasing proportions of false and missed detections with increasing data set duration, due to the higher proportion of acoustic noise and lower overall call rates in longer recordings. However, the automated detector missed only one calling “bout” (2 or more calls within a 10-minute span) of the 18 bouts present in the week-long recording. Despite the high number of false detections and missed individual calls, spectrogram cross-correlation was useful to guide a human analyst to sections of data with potential right whale calling bouts. Upon reviewing automatic detection events, the analyst could quickly dismiss false detections and search recordings before and after correct detections to find missed calls, thus improving the efficiency of searching for a small number of calls in long-term (months- to years-long) recordings.

1. INTRODUCTION

Long-term, passive acoustic recorders are useful tools for monitoring some marine mammal populations, with potential applications ranging from providing information on behavioral ecology and abundance, to near-real-time localization and tracking of calling animals (*e.g.* Thompson and Friedl 1982, Clark *et al.* 1996, Stafford *et al.* 2001, Gillespie and Leaper 2001, Moscrop *et al.* 2004, Mellinger *et al.* 2004a, b). We used autonomous, bottom-mounted Acoustic Recording Packages (ARPs) (Wiggins 2003) to provide long-term recordings of critically endangered North Pacific right whales (*Eubalaena japonica*) (Brownell *et al.* 2001) and other baleen whale species in the southeast Bering Sea. Here, small numbers (tens) of right whales have been regularly observed since 1996 in the middle-shelf region (between the 50 m and 100 m isobaths) in summer months (Goddard and Rugh 1998, LeDuc *et al.* 2001, LeDuc 2004). We deployed and recovered five ARPs from four sites in the right whale sighting region in 2000-2002 (Figure 1). The ARPs recorded sound continuously in a frequency range (5 to 250 Hz) encompassing that of most North Pacific right whale calls (McDonald and Moore 2002), and provided 36

instrument-months totaling over 100 gigabytes of data.

Because of a paucity of data on eastern North Pacific right whales, each recorded right whale call could contribute to a better understanding of this population. However, manually detecting each right whale call in this large data set would potentially require hundreds of hours of human effort to scan spectrograms visually and to listen to recordings. In contrast, a computer using automated detection software could potentially process a year-long data set within hours to days, and human effort could be focused on reviewing automated detection results and searching for additional calls near times of automatic call detections. We found that automated call detection using spectrogram cross-correlation was effective for detecting bouts of right whale calling in long-term acoustic recordings from the Bering Sea. This paper evaluates the performance of spectrogram cross-correlation in detecting right whale calls within a subset of Bering Sea ARP data.

North Pacific right whales were first recorded in the Bering Sea in 1999 by McDonald and Moore (2002). The most common right whale call type (85%, $n=511$) was an ‘up’ call, sweeping up in frequency on average from 90 to 150 Hz in 0.7 s (McDonald and Moore 2002). North Pacific right whale calls and the proportion of different call types

were similar to call repertoires of other right whale species (*Eubalaena* spp.) (Clark 1982, Matthews *et al.* 2001). Similarly to North Atlantic right whale calls (Matthews *et al.* 2001, Vanderlaan *et al.* 2003), North Pacific right whale calls were clustered in sporadic ‘bouts’ lasting several minutes, with longer silences (tens of minutes to hours) between bouts (McDonald and Moore 2002).

Automated right whale call detection in Bering Sea acoustic data was challenging for a number of reasons. Right whale call durations were brief (≤ 1 s), and calls were variable in duration, start and end frequencies, and frequency sweep rates (Figure 2, McDonald and Moore 2002). Calls may also have become distorted at the receiver due to the dispersion of normal modes over the flat, shallow continental shelf (Wiggins *et al.* 2004). Overall, calls received on ARPs were infrequent and the total number of calls was low. Flow and strum noise on hydrophones was frequently exacerbated by storms and strong tidal currents characteristic of the Bering Sea middle-shelf (Bond and Adams 2002, Coachman 1986). Also adding to the challenge of automated detection, humpback whales (*Megaptera novaeangliae*) produced sounds (Figure 3), including upswept calls, in the same frequency band used by right whales and recorded by the ARPs.

A variety of automated call detection techniques are available, including matched filtering, spectrogram correlation, energy summation, and neural networks (Stafford *et al.* 1998, Mellinger and Clark 2000, Mellinger 2004, Mellinger *et al.* 2004b). The performance of each of these techniques often depends on the characteristics of a particular species’ acoustic repertoire and behavior and the physical environment in which they are recorded. For example, matched filtering works well when calls are highly stereotyped, and energy summation

works well for species that call often and in a frequency band isolated from other sounds (*e.g.*, calls from other species, ship engine noise, cable strumming) (Mellinger 2004). Some techniques, such as neural networks, require a large training set of calls. Right whale calls in Bering Sea ARP recordings were not well suited to matched filtering, energy summation, or neural networks.

The challenges of right whale call detection in our data set led us to investigate spectrogram cross-correlation with a synthetic kernel because this method a) does not require a

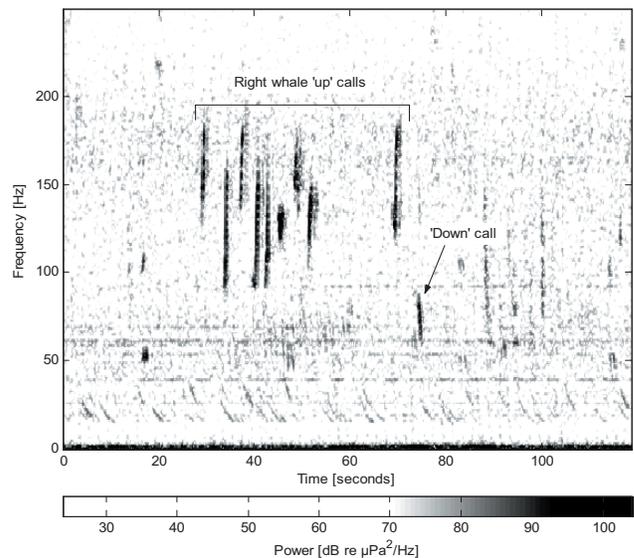


Figure 2. Two-minute excerpt of hour-long recording containing right whale calls, recorded by ARP at site C, 3 October 2000. Spectrogram parameters: 512 point frame and FFT length with same size Hanning window, 75% overlap, for a filter bandwidth of 4.0 Hz. Also visible throughout are fin whale downsweeps from 35-15 Hz.

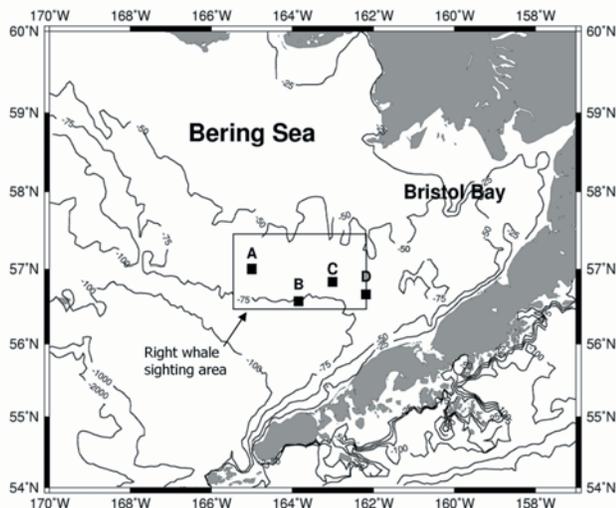


Figure 1. ARP sites A through D, year 2000-2002. Recordings in 2000-01 were from all four sites; recordings the following year (2001-02) were from site C only. Bathymetric contours are displayed at 25-meter increments for depths up to 100 m, and at 1000-m increments for depths of more than 100 m. Right whale visual sighting locations in the Bering Sea since 1996 are bounded by the ‘sighting area’ box.

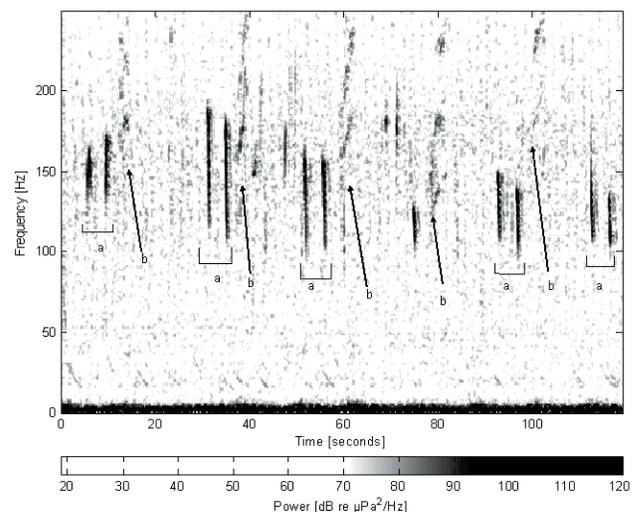


Figure 3. Two-minute excerpt containing humpback whale calls from week-long sound recording, recorded by ARP at site C. Repeated calls are labeled ‘a’ and ‘b’ to show pattern. Spectrogram parameters same as in Figure 2.

large training set of calls, b) may be more suited to detecting brief and infrequent calls in a large and often noisy data set, and c) may be less sensitive to variation and distortion among calls than the other techniques (Mellinger 2004). We configured the automated detector to detect 'up' calls, the most common call recorded from North Pacific right whales in the Bering Sea (McDonald and Moore 2002, Munger and Sauter unpub. data, Munger and Rankin unpub. data). In this study, we optimized automated detection parameters using short-duration recordings, and then evaluated the performance of optimized parameters in processing a week-long recording.

2. METHODS

2.1. Data sets

ARPs were configured to record continuously at a sampling frequency of 500 Hz, with a frequency response of -152 dB re 1 V/ μ Pa, flat within 1 dB over the 5-250 Hz frequency band (Wiggins 2003). Acoustic data were digitized to 16-bit samples and stored on computer hard disks to be analyzed after instrument recovery.

Three subsets of acoustic data from the year-2000 ARP recordings were used to test the automated detector: short, intermediate, and long duration recordings. Short and intermediate duration recordings were chosen from previously manually processed data to provide the detector with training sets of calls with which to optimize detection parameters. Short recordings contained right whale 'up' calls, humpback whale calls, a combination of both, or no discernible calls; intermediate and long recordings were continuous sections of data containing periods of noise and calls from right and humpback whales.

The first data subset consisted of twelve short (1-to 5-minute) recordings made at different times by four ARPs during October-December 2000. Six of these recordings contained calls in the 80-250 Hz bandwidth, with a total of 26 right whale 'up' calls. The other six recordings contained no calls in that frequency range and varied in acoustic noise levels (Figure 4). The average overall right whale call rate in the short data set was 1.96 calls per minute.

The second data subset consisted of four intermediate-length recordings, 65 minutes each (Figure 2), recorded simultaneously on each of the four ARPs on October 3, 2000. Each of these recordings contained right whale and humpback whale calls in the 80-250 Hz band, including 72 right whale 'up' calls. The overall average right whale call rate in intermediate-length recordings was 0.28 calls per minute. The intermediate-length recordings did not have any data in common with the short recordings.

The third data subset was a single recording approximately one week in length, taken from the ARP at site C on 2-9 October 2000. This long recording did not share common data with the short recordings, but did encompass the hour recorded by ARP C in the intermediate-length data set. Whale calls were present in at least five days of the week-long data

set; these included humpback calls (Figure 3) and 146 right whale 'up' calls, the majority of which were recorded during the first three days. The average right whale call rate over the week-long recording was 0.015 calls per minute.

2.2. Manual call detection

After evaluating right whale acoustics literature and discussing right whale call types with colleagues, the human analyst (LMM) visually scanned spectrograms and listened to potential calls throughout the first month of ARP recordings (October 2000). One difficulty in detecting right whale calls was distinguishing between calls of humpback whales and right whales. Humpback whales produced some sounds in the same frequency band as right whales, including short-duration upswept or downswept calls. The most important

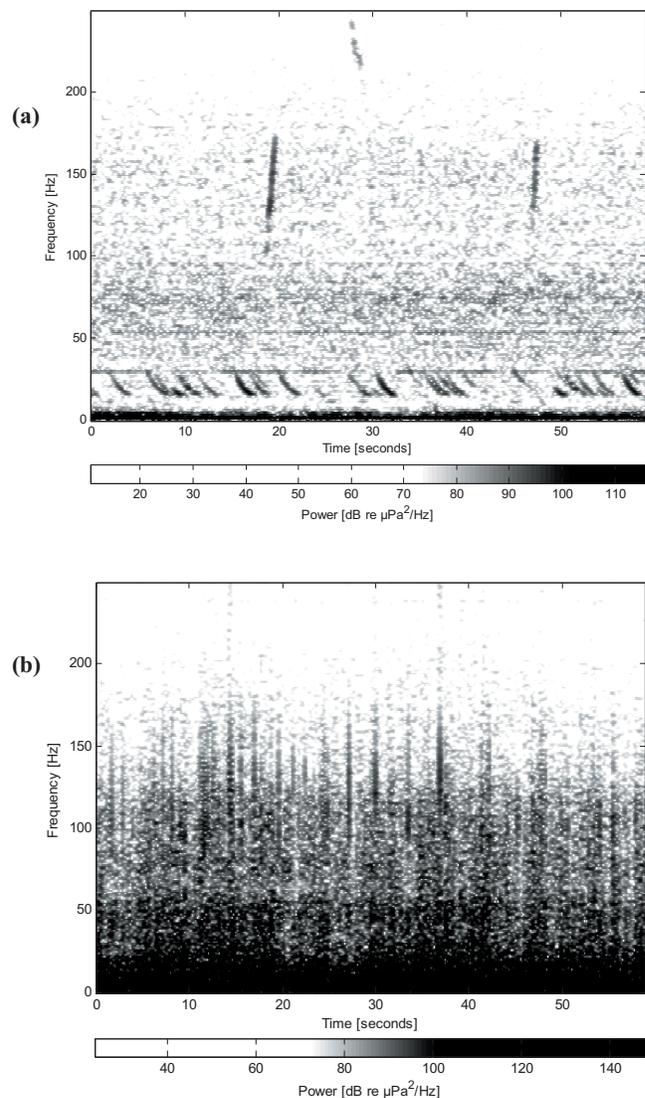


Figure 4. Two examples of sixty-second recordings containing a) right whale 'up' calls at approximately 18 s and 47 s, and b) only noise. Spectrograms parameters same as in Figure 2 except 90% overlap.

distinguishing feature in our Bering Sea data set proved to be the temporal pattern of calls. Right whales produced calls (Figure 2) in sporadic bouts, whereas humpback whales produced calls in consistent, repeated patterns (Figure 3). Patterned humpback calling (song) has been reported in late summer/early fall on other northern feeding grounds as well (Mattila *et al.* 1987, McSweeney *et al.* 1989). In addition, ARP recordings of humpback calls and call series often contained harmonics and higher-frequency components, whereas right whale calls were typically tonal upsweeps without harmonics.

The human analyst (LMM) used a software program (*Triton*, Wiggins 2003) written in MATLAB® (The MathWorks, <http://www.mathworks.com>) to generate and display spectrograms of the ARP data sets. Time series were transformed into the frequency domain using a Fast Fourier Transform (FFT) with a Hanning window (Oppenheim and Schaffer 1999). FFT and window length were both 512 points (1.024 s) and overlap was 75-90%. Graphical gain and contrast were adjusted to give the best resolution of the spectrogram. During visual scanning of sequential spectrograms, the time-frequency display window was 0-250 Hz in frequency and usually 60-120 s in duration. When the analyst detected a potential right whale call in the displayed spectrogram, the call portion of the display was expanded in time and spectral parameters were adjusted to ‘sharpen’ the image--for example, by reducing FFT and window length and increasing the amount of overlap. In addition, potential right whale calls were also played on speakers, to provide the analyst an opportunity to aurally detect and distinguish right whale calls from humpbacks if visual detection was ambiguous.

The analyst noted only right whale ‘up’ calls for the purposes of this comparative study because the automated detector was configured to detect only ‘up’ calls. The set of manually-picked right whale ‘up’ calls provided the basis for comparing automated detection results.

2.3. Automated detection

We used the software program *Ishmael* (Mellinger 2001) for call detection by spectrogram cross-correlation. Spectrograms were generated in *Ishmael* using the same parameters as used in *Triton* to manually detect calls: frame, FFT, and Hanning window length were equal to 512 points (1.024 s), and overlap was 75-94%. ARP spectral data were cross-correlated with a synthetic spectrogram kernel (Mellinger and Clark 2000), which we based on the ‘up’ calls found in our data sets and consistent with those described in McDonald and Moore (2002). The synthetic call kernel consisted of piecewise, continuous line segment(s) defined by start and end times and their corresponding start and end frequencies (Figure 5). Other detection parameters that were adjusted included the instantaneous bandwidth of the synthetic call kernel (Figure 5), detection threshold, minimum and maximum duration above the detection threshold, and

spectrogram equalization time constant (time-averaging to smooth out background noise) (Van Trees 1968; Mellinger 2001, 2004; Mellinger *et al.* 2004a). The minimum time between detections was set to 0 seconds to avoid missing close or overlapping calls.

The spectrogram cross-correlation output is a time series of the unnormalized cross-correlation, which varies with the closeness of the match between the data and the predefined kernel; function peaks above a user-specified threshold are counted as detection events. If the parameters we chose resulted in zero detection events, we discarded that set of parameters and did not include them in this analysis. When detection events occurred, we adjusted one parameter at a time and observed the resulting effect on detector performance. If performance improved and resulted in fewer false detections and/or missed detections, we adjusted the other parameters in an attempt to further minimize missed detections and false detections.

We ran 62 automated detection trials using the short recordings, 22 trials using the intermediate-length recordings, and 1 trial using the week-long recording. Each automated detection event was saved individually as a short (~10 s) sound file. After each detector trial, a human analyst examined each individual detection event to verify whether the detection was correct. Automated detections were classified as correct detections (right whale ‘up’ calls) or false detections. False detections were further categorized as other biological sounds, including non-upswept calls or calls identified to be from humpback whales, or noise, in which no call was present.

We compared the performance of various detection parameters by plotting receiver operating characteristic (ROC)

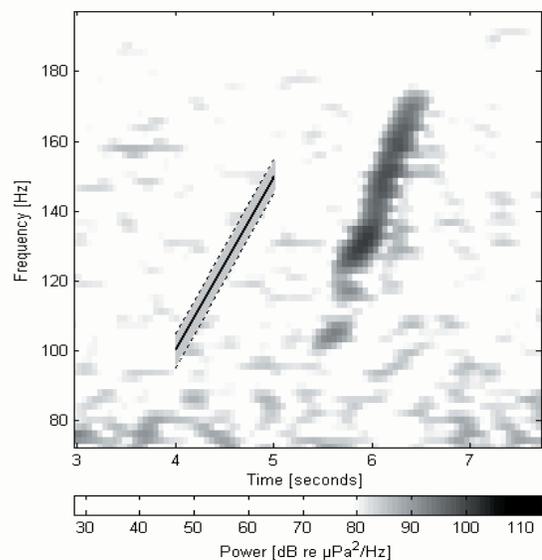


Figure 5. Synthetic call kernel (light gray) from 100 to 150 Hz, lasting 1 s, with ‘kernel bandwidth’ of 10 Hz bounded by dotted lines. The synthetic kernel precedes a spectrogram of a right whale call, enlarged from Figure 4a at 18 s. Spectrogram parameters: 256-point frame, FFT, and Hanning window length, and 90% overlap, for a filter bandwidth of 7.9 Hz.

curves illustrating the trade-off between false detections and missed calls. False detections were expressed as a percentage of the total number of automated detections. Missed calls were expressed as percentages of the total number of ‘up’ calls in the data set, which was defined as the number of manually detected ‘up’ calls. We designated an acceptable missed call threshold of 20%, and defined ‘optimal’ detection parameters as those that minimized false detections while missing fewer than 20% of calls. We set this missed call threshold because right whale calls were rare in our data set and we wished to detect a substantial majority of them; although this caused an increase in false detections, reviewing and discarding false detections was still a simple and relatively fast process for an analyst compared to thoroughly manually processing the entire recording. Optimal parameters from trials using short recordings were included in detector trials run on intermediate-length recordings, and the optimal parameters from trials using intermediate recordings were used to process the week-long recording.

3. RESULTS

For ease of interpretation, we separated automated detection results for short (minutes-long) recordings into results using a single synthetic kernel, and those using different kernel types with varying slopes and numbers of segments. The automated detector that performed best on the short recordings used a 1 s, 100-150 Hz synthetic kernel (Table 1, Figure 6). The optimal parameters (resulting in fewer than 20% missed calls and minimal false detections) with this single kernel type resulted in 19% missed detections (5 of 26 calls) and 25% false detections (7 of 28 total detections) (Figure 6, Table 2). A large proportion (86%) of false detections contained other biological sounds.

Table 3 shows the detection parameters used with synthetic kernels that consisted of one or more segments of varying duration and start/end frequencies. The corresponding ROC curves for those parameters were plotted in Figure 7. The optimal synthetic kernel in this case (resulting in fewer than 20% missed calls and minimal false detections), consisted of 2 segments: the first 1 s and 100 to 150 Hz, and the second 0.5 s from 150 to 180 Hz. These parameters resulted in 19% missed detections and 42% false detections (Figure 7, Table 4), 80% of which were other biological sounds. The varied synthetic call kernels that we tested did not perform as well as the single-segment 1 s, 100-150 Hz kernel.

Detection parameters and results for intermediate-length recordings are shown in Table 5, Figure 8, and Table 6. In addition to varying the same parameters as in short-recording trials, we used spectrogram equalization (time-averaging to smooth background noise) in some trials; this was not done for short recordings because averaging over seconds was inappropriate for recordings lasting tens of seconds. For the same acceptable level of missed detections (20%), the optimal detection parameters resulted in 69% false detections and 19% missed calls. 31% of these false detections in

intermediate-length recordings were other biological sounds. (Figure 8, Table 6). Although we tested varying synthetic kernels, the optimal detection parameters were again based on a single 1 s, 100-150 Hz segment, and did not employ spectrogram equalization.

The detection parameters used for the week-long recording were the optimal parameters resulting from trials using hour-long recordings. The detection results using the week-long recording are summarized in Table 7 and displayed as a single data point on Figure 8a. False detections comprised 98% of the total number of detections, and approximately 38% of detectable calls were missed. Of the false detections, 10% were other biological sounds. Figure 9 compares the number of manual and automated detections of right whale ‘up’ calls over the first three days (when most of the right whale calls were detected) of the week-long ARP recording. We defined a calling ‘bout’ as at least 2 calls within a ten-minute time span; although not all calls in a single bout were detected, the automated detector missed only one of 18 total bouts in the week-long recording, and missed 3 calls occurring singly.

To investigate whether false detection rates were related to acoustic noise levels, we compared noise levels in the recordings by calculating average spectral levels between

Symbols in Figure 6:	◆	■	◇	□	△	×	+
Kernel width (Hz)	10	10	10	10	10	7-12	9-14
Detection threshold	5	6	3-7	2-8	3-5	4	5
Minimum duration above threshold (s)	0.2-0.7	0.3-0.5	0.4	0.5	0.6	0.5	0.5
Maximum duration above threshold (s)	2	2	2	2	2	2	2

Table 1. Detection parameters tested in short recording trials; the synthetic call kernel in all of these tests was a 1 s, 100-150 Hz line segment as in Figure 5. The ranges of varying parameters are shown in bold type. Symbols correspond to markers in Figure 6.

100 and 150 Hz over 1-minute time intervals in hour-long recordings and 10-minute intervals in the week-long recording (time intervals were shorter for the hour-long recordings to give better graphical resolution). The percentage of false detections during the intermediate-length test recordings differed significantly between each of the four ARPs, and was highest on ARP A (average false detection proportions: A=89.8%, B=57.0%, C= 59.8%, D= 77.3%; ANOVA, $p<0.05$), which also had the highest average noise level over the hour of recording (Figure 10).

Noise levels varied more over one week on an individual instrument (ARP C) than they did between instruments during the hour recorded in the intermediate-length test data (Figure 10). During the week recorded on ARP C, a semidiurnal tidal signature was apparent during the first three days of recording, and an overall rise in noise during days 280-282 was caused by a storm. Peaks in noise on days 279 and 280 were related to passing ships—closer inspection of spectrograms revealed long, continuous tones at 60 Hz and higher harmonics typical of engine-related noise.

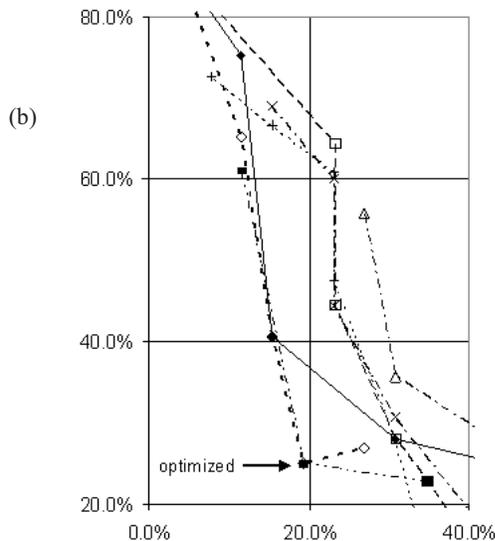
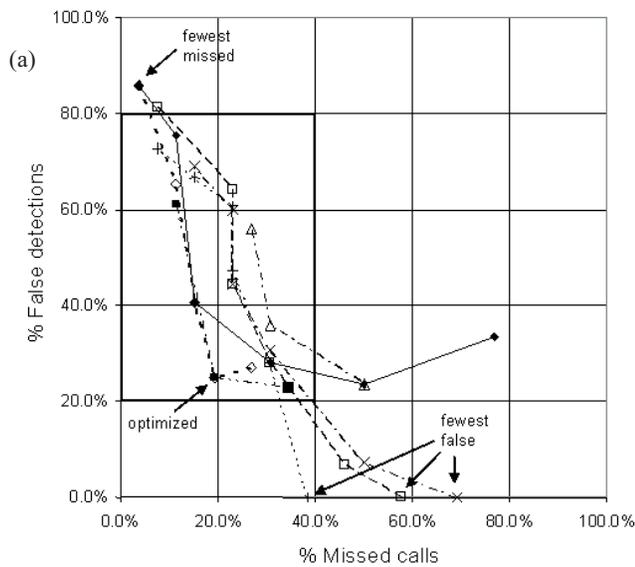


Figure 6. a) Results of automated detection trials using short recordings, and parameters and symbols from Table 1. Each curve is the result of varying one detection parameter and measuring resulting rates of false detections and missed calls. False detections are expressed as a percentage of the total number of automated detections; missed calls are expressed as a percentage of the total number of manually detected calls. Results for 'optimized', 'fewest missed', and 'fewest false' data points are given in Table 2. b) Area within thickened line in 6a is expanded in Figure 6b.

4. DISCUSSION

Using spectrogram correlation with manually optimized detection parameters, the automated detectors we tested performed best on the short (minutes-long) sound recordings, with increasing proportions of false and missed detections as the recording duration increased. The increase in the proportion of false detections with the recording length was expected, because the longer recordings contained longer periods of noise relative to the number of right whale calls

	Optimized (<20% missed)	Fewest missed	Fewest false (38% missed, 0% false)
Kernel width (Hz)	10	10	10
Detection threshold	6	3	3-7
Minimum duration above threshold (s)	0.4	0.4	0.4
Missed detections out of 26 calls	5	1	10
False detections: total	7	152	0
False detections: other call types/species	(6)	(61)	(0)
Total number of detections	28	177	16

Table 2. Detection parameters and results for the 'optimized' data point (using a predefined threshold of 20% missed calls), 'fewest missed' calls, and 'fewest false' detections (Figure 6) using short recordings and synthetic kernel of 1 s, 100-150 Hz. False detection total includes other biological sounds, which are reported in parentheses.

present and provided more opportunities for the detector to produce false detections. The short- and intermediate-length recordings were used to optimize detection parameters and consequently represented relatively high rates of right whale calling (approximately 2 calls/minute in the minute(s)-long recordings and 0.3 calls/min in the hour-long recordings),

Symbols in Figure 7: ◆ ■ ▲ ◇ □ △ × + - ○

Number of segments	1	1	1	1	1	1	1	3	2	1
Segment length (s)	1	0.5	0.5	0.5	0.6	0.5-1	1	0.3	0.5	0.7
Sweep rate (Hz/s)	60	30	30	30	42	42-50	60	37	70	86
Start Frequency (Hz)	100	100	125	110-125	100	100	90	100	90	90
End frequency (Hz)	160	125	150	135-150	125	150-180	150	151	160	150
Kernel width (Hz)	9-11	8-10	8-10	9-10	8-9	8-10	9-10	10	10	10
Detection threshold	4-5	3-4	3	3-3.5	3-4	3-4	4	3	3	4
Minimum duration above threshold (s)	0.4-0.5	0.5	0.5	0.4-0.5	0.5	0.5	0.5	0.5	0.5	0.5
Maximum duration above threshold (s)	2	2	2	2	2	2	2	2	2	2

Table 3. Range of detection parameters tested in short recording trials when varying the structure of the synthetic call kernel. The ranges of varying parameters are shown in bold type. Symbols correspond to markers in Figure 7.

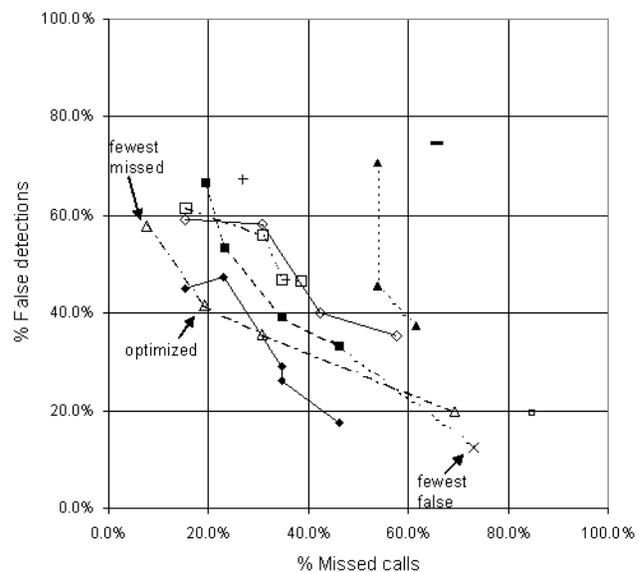


Figure 7. Results of automated detection trials using short recordings and varying the synthetic kernel structure. Symbols as in Table 3, and terminology in Figure 6. Some of the 'curves' here consist of a single point. Results for 'optimized', 'fewest missed', and 'fewest false' data points are given in Table 4.

whereas the week-long recording contained 0.02 calls/min. Because the longer data sets contained more calls in total (72 calls in intermediate-length data set and 146 calls in week-long data set), there was increased potential for variation among calls, possibly contributing to the higher proportion of missed detections caused by a mismatch between calls and the synthetic kernel.

The automated detection trials also resulted in high rates of detection of other biological sounds: over 80%, 31%, and 10% of the false detections were biological sounds for short, intermediate, and long recordings, respectively. These other sounds included upswept calls from humpback whales, as well as other call types (down-swept calls, moans, pulses) that potentially could have been made by humpbacks or right whales. Although these were classified as false detections (because they were not right whale 'up' calls), a human analyst in practice would likely be interested in reviewing these sounds as well, especially if the goal is to correctly detect and classify each rare right whale call during the post-processing of a large data set.

The automated detector produced over 90% false detections and missed over one-third of the right whale 'up' calls in the week-long data recording. These results were poor compared to some other marine mammal acoustic detection studies, in which automated detection software missed relatively fewer calls, categorized a greater proportion of calls correctly, and produced a smaller percentage of false detections (Mellinger and Clark 2000; Niezrecki *et al.* 2003; Mellinger *et al.* 2004a,b). Some factors contributing to the high missed call rate in our study were variability among calls (McDonald and Moore 2002), distortion resulting from waveform dispersion (Wiggins *et al.* 2004), and high acoustic noise levels resulting in decreased signal-to-noise ratios (SNR) of calls.

Acoustic noise recorded by hydrophones in the Bering Sea was often high due to strong tidal currents and frequent storms (Figure 10b). The lack of detections in the week-long recording during approximately days 280-282, when the noise level was highest (Figure 10), could have been due to masking by that noise or to an actual lack of whale calls. A rise in noise level may decrease the acoustic detection range and could explain the lack of detected calls. High

	Optimized (<20% missed)	Fewest missed	Fewest false
Number of segments	2	2	1
Segment durations (s)	1, 0.5	0.6, 0.5	1
Start-middle-end frequency (Hz)	100-150-180	100-125-150	90-150
Kernel width (Hz)	10	9	9
Detection threshold	4	3	4
Minimum duration above threshold (s)	0.5	0.5	0.5
Missed detections out of 26 calls	5	2	19
False detections: total	15	33	1
False detections: other call types/species	(12)	(17)	(1)
Total number of detections	36	57	8

Table 4. Detection parameters and results for the 'optimized' data point (using predefined threshold of 20% missed calls), 'fewest missed' calls, and 'fewest false' detections (Figure 7) using short recordings and varying the synthetic kernel structure. False detection total includes other biological sounds, which are reported in parentheses.

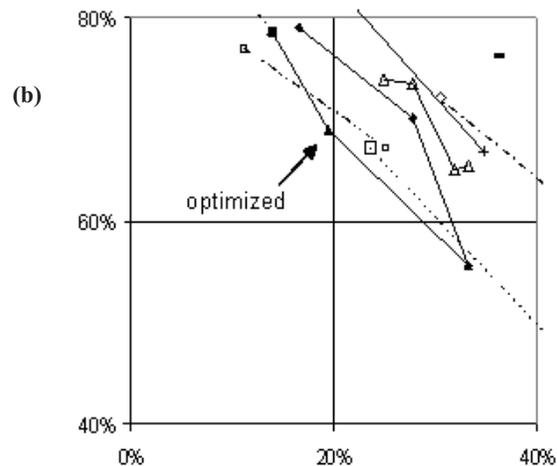
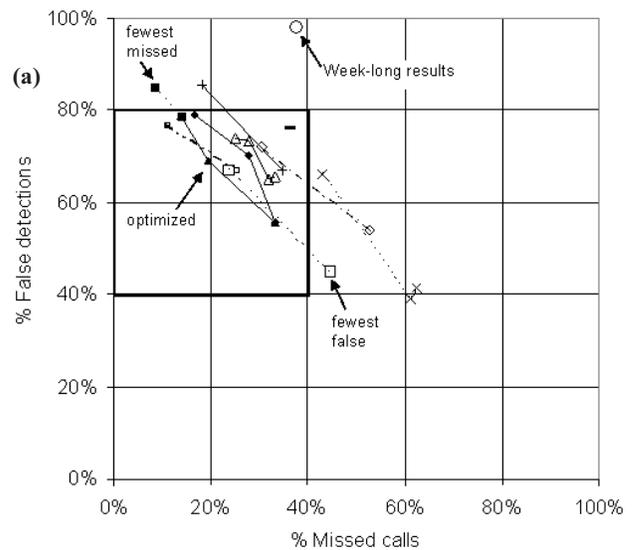


Figure 8. a) Results of automated detection trials using intermediate-length (symbols in Table 5) and week-long (marked by single open circle) recordings. Terminology same as in Figure 6. Some of the 'curves' here consist of a single point. Results for intermediate-length points labeled 'optimized', 'fewest missed', and 'fewest false' are given in Table 6. Results from week-long data set (open circle) given in Table 7. b) Area within thickened line in 8a is expanded in Figure 8b.

Symbols in Figure 8:

	◆	■	▲	◇	□	△	×	+	-	○
Number of segments	1	1	1	1	1	1	1	1	2	1
Segment duration (s)	1	1	1	1	1	1	1	1.1	1, 0.5	1
Sweep rate (Hz/s)	50	50	50	50	50	50	50	45-55	50, 60	55-60
Start Frequency (Hz)	100	100	100	100	100	100	100	100	100	100
End frequency (Hz)	150	150	150	150	150	150	150	150	180	155-160
Kernel width (Hz)	10	11	10-11	9-10	10-11	10-11	10	10	10	11
Detection threshold	5-6	5.5-6	6	4	5	5-6	4.5	5	4	5.5
Minimum duration above threshold (s)	0.4	0.4	0.4	0.5	0.5	0.4	0.5	0.4	0.5	0.4
Maximum duration above threshold (s)	2	1.5-2	2	2	2	2	2	2	2	2
Equalization constant (s)	None	None	None	None	None	3-5	2-3	None	None	None

Table 5. Detection parameters tested in intermediate-length recording trials. The ranges of varying parameters are shown in bold type. Symbols correspond to markers in Figure 8.

acoustic noise levels also contributed to high false detection rates; during the hour-long recording, ARP A had both the highest false detection rate and highest average noise level in the call frequency band (Figure 10). It is not clear whether the relatively higher noise on ARP A was due to differences in instrument calibration or actual differences in acoustic

	Optimized ($<20\%$ missed)	Fewest missed	Fewest false
Kernel width (Hz)	10.5	11	10
Detection threshold	6	5.5	5
Minimum duration above threshold (s)	0.4	0.4	0.5
Maximum duration above threshold (s)	2	1.5	2
Equalization time constant	None	None	3
Missed detections out of 72 calls	14	6	44
False detections: total	129	376	18
False detections: other call types/species	(40)	(88)	(12)
Total number of detections	187	376	46

Table 6. Detection parameters and results for the ‘optimized’ data point (using predefined threshold of 20% missed calls), ‘fewest missed’ calls, and ‘fewest false’ detections (Figure 8) using intermediate-length recordings. The synthetic kernel in all three cases was 1 s, 100-150 Hz, although alternative synthetic kernels were also tested. False detection total includes other biological sounds, which are reported in parentheses.

Results for week-long recording:

Missed detections out of 146 calls	55
False detections: total	4566
False detections: other call types/species	(458)
Total number of detections	4657

Table 7. Detection results for week-long recording. Detection parameters are same as ‘optimized’ parameters from trials using intermediate-length recordings (Table 6).

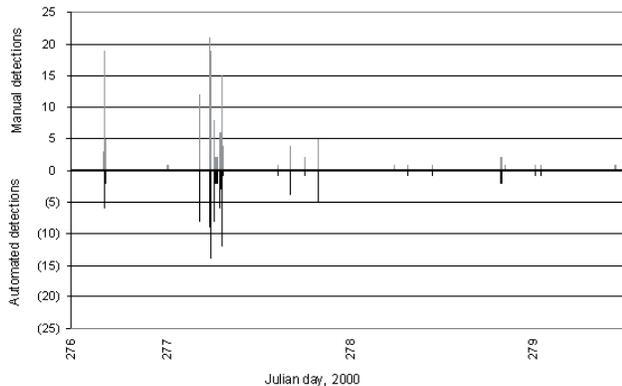


Figure 9. Manual detections per ten-minute time bin (gray bars, upper half of plot) in the first three days of the week-long recording, and automated detections (black bars, lower half of plot) within the same recording, using ‘optimized’ parameters from intermediate-length recording trials. In this recording, one ‘bout’ of calls, defined as at least two calls per ten-minute span, was missed on day 276, out of 18 bouts total within the entire week. Four single calls were also missed by the detector on days 277-279. On day 282 (not pictured), two calls were detected manually, one of which was detected automatically.

noise. Noise levels varied more over a long duration of time on a single instrument than between instruments during the same short time period (Figure 10); therefore any effects of ARP calibration on detector performance were probably overshadowed by the much larger fluctuations in acoustic noise over time due to events such as tides, storms, and passing ships.

Despite the high rates of false and missed detections, automated call detection by spectrogram correlation was nevertheless useful for our complete ARP data set. Although a human analyst reviewed all of the automated detections, this process was considerably more time-efficient than thoroughly scanning the entire data set manually. In our trial using the week-long recording, the detection parameters we used missed only one of 18 right whale calling ‘bouts’ (Figure 9). Automated spectrogram correlation, optimized for a low number of missed detections, was thus helpful in directing a human analyst to periods in the data when additional calls

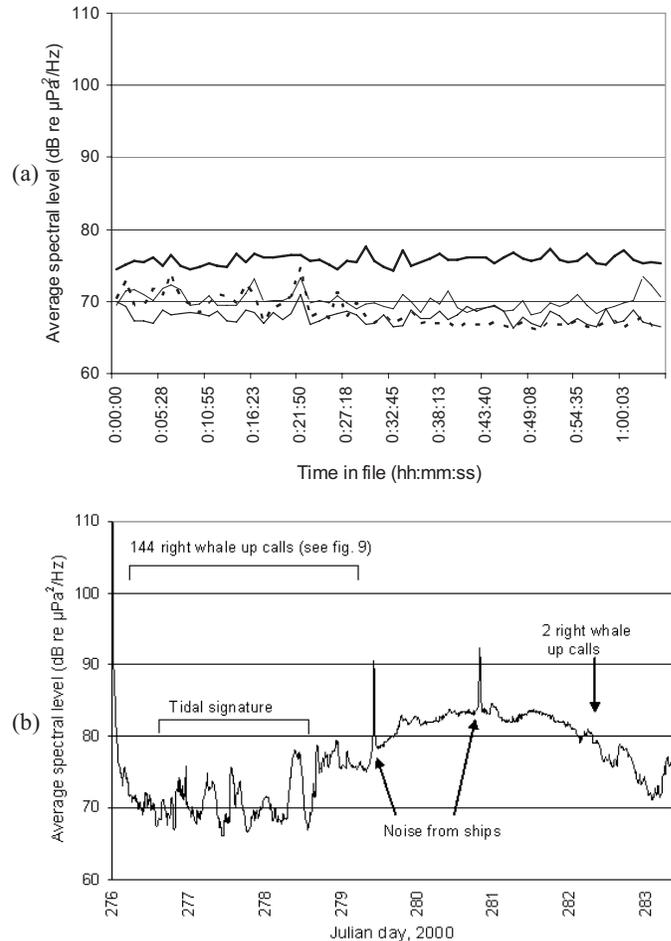


Figure 10. a) Noise levels between 100-150 Hz on each ARP during intermediate-length (65-minute) recordings, recorded simultaneously on each instrument. b) Noise levels between 100-150 Hz on ARP C during week-long recording. Tidal signature and spikes in noise from ships are labeled, as are days of right whale calls. Overall increase in noise on days 280-282 due to storm. Intermediate-length recordings taken from day 277, hour 0400-0505.

were likely to be found near the automated call detection. The combination of automated detection with manual verification and focused searching has been used effectively in detecting North Pacific right whale calls in the Gulf of Alaska (Waite *et al.* 2003, Mellinger *et al.* 2004b), as well as in other long-term right whale data sets (Clark *et al.* 2000).

Due to the paucity of data on right whales in the eastern North Pacific, our primary goal in developing an automated right whale call detector was to maximize the number of right whale calls detected, and the concomitant increase in high false detection rates was acceptable during this study. Detection techniques other than spectrogram cross-correlation, such as neural networks (Mellinger 2004), may become more feasible as we increase the set of known calls recorded in the presence of North Pacific right whales. Current and future deployments of passive acoustic recorders in the Gulf of Alaska and Bering Sea will provide new data that will require efficient processing and benefit from improved automated detection techniques. For the ARP data set described in this study, automated detection using spectrogram correlation was useful to direct a human analyst to potential right whale calling bouts and was more time-efficient than manual call detection.

ACKNOWLEDGEMENTS

The authors thank NOAA/National Marine Mammal Laboratory for supporting the fabrication, deployment, and recovery of Acoustic Recording Packages in the Bering Sea (JIMO budget #NA17RJ1231). We also thank the North Pacific Marine Research Institute (project #T-2100) and North Pacific Research Board (project #R0307) for supporting data processing, outreach and travel to study areas. The work was also supported by ONR contracts N00014-03-1-0099 and N00014-03-1-0735. This is PMEL contribution #2700.

REFERENCES

- Bond, N.A. and J.M. Adams. 2002. Atmospheric forcing of the southeast Bering Sea Shelf during 1995–99 in the context of a 40-year historical record. *Deep-Sea Res. II* 49: 5869–5887.
- Brownell, R.L., P.J. Clapham, T. Miyashita, and T. Kasuya. 2001. Conservation status of North Pacific right whales. *J. Cetacean Res. Manage.* (Special Issue) 2:269-286.
- Clark, C. W., R. Charif, S. Mitchell, and J. Colby. 1996. Distribution and behavior of the bowhead whale, *Balaena mysticetus*, based on analysis of acoustic data collected during the 1993 spring migration off Point Barrow, Alaska. *Rep. Int. Whal. Commn.* 46: 541-552.
- Clark, C.W. 1982. The acoustic repertoire of the southern right whales, a quantitative analysis. *Anim. Behav.* 30: 1060-1071.
- Clark, C.W., D. Gillespie, A. Moscrop, T. Fowler, T. Calupca, and M. Fowler. 2000. Acoustic sampling of right whale vocalizations in the Great South Channel using sea-floor pop-up recorders. Report of 33 - Vol. 33 No. 2 (2005)

the Right Whale Consortium, Boston, Massachusetts.

- Coachman, L.K. 1986. Circulation, water masses, and fluxes on the southeastern Bering Sea shelf. *Continental Shelf Res.* 5: 23-108.
- Gillespie, D. and R. Leaper (eds.). 2001. Report of the Workshop on Right Whale Acoustics: Practical Applications in Conservation. International Fund for Animal Welfare, Yarmouth Port, MA 02675, USA. 23 pp.
- Goddard, P.D., and D.J. Rugh. 1998. A group of right whales seen in the Bering Sea in July 1996. *Marine Mamm. Sci.* 14: 344-349.
- LeDuc, R.G. 2004. Report of the results of the 2002 survey for North Pacific right whales. NOAA Technical Memorandum NMFS-SWFSC-357, NMFS Southwest Fisheries Science Center, La Jolla.
- LeDuc, R.G., W.L. Perryman, J.W. Gilpatrick, Jr., J. Hyde, C. Stinchcomb, J.V. Carretta, and R.L. Brownell, Jr. 2001. A note on recent surveys for right whales in the southeastern Bering Sea. *J. Cetacean Res. Manage.* (Special Issue) 2: 287-289.
- Matthews, J.N., S. Brown, D. Gillespie, M. Johnson, R. McLanaghan, A. Moscrop, D. Nowacek, R. Leaper, T. Lewis, and P. Tyack. 2001. Vocalisation rates of the North Atlantic right whale (*Eubalaena glacialis*). *J. Cetacean Res. Manage.* 3: 271-282.
- Mattila, D. K., L. N. Guinee, and C. A. Mayo. 1987. Humpback whale songs on a North Atlantic feeding ground. *J. Mamm.* 684: 880–883.
- McDonald, M.A., and S.E. Moore. 2002. Calls recorded from North Pacific right whales (*Eubalaena japonica*) in the eastern Bering Sea. *J. Cetacean Res. Manage.* 4: 261-266.
- McDonald, M.A. 2004. DIFAR hydrophones applied to whale research. *Canad. Acoust.* 32: 155-160.
- McSweeney, D. J., K. C. Chu, W. F. Dolphin, and L.N. Guinee. 1989. North Pacific humpback whale songs: a comparison of southeast Alaskan feeding ground songs with Hawaiian wintering ground songs. *Marine Mamm. Sci.* 5: 139–148.
- Mellinger, D.K. 2001. Ishmael 1.0 User's Guide. NOAA Technical Report OAR-PMEL-120, NOAA Pacific Marine Environmental Laboratory, Seattle. Also available online at <http://cetus.pmel.noaa.gov/cgi-bin/MobySoft.pl>. 30 pp.
- Mellinger, D.K. 2004. A comparison of methods for detecting right whale calls. *Canad. Acoust.* 32: 55-65.
- Mellinger, D.K., and C.W. Clark. 2000. Recognizing transient low-frequency whale sounds by spectrogram correlation. *J. Acoust. Soc. Am.* 107: 3518-3529.
- Mellinger, D.K., K.M. Stafford, and C.G. Fox. 2004a. Seasonal occurrence of sperm whale (*Physeter macrocephalus*) sounds in the Gulf of Alaska, 1999-2001. *Marine Mamm. Sci.* 20: 48-62.
- Mellinger, D.K., K.M. Stafford, S.E. Moore, L. Munger, and C.G.

Fox. 2004b. Detection of North Pacific right whale (*Eubalaena japonica*) calls in the Gulf of Alaska. *Marine Mamm. Sci.* 20: 872-879.

Moscrop, A., J. Matthews, D.Gillespie, and R. Leaper. 2004. Development of passive acoustic monitoring systems for northern right whales. *Canad. Acoust.* 32: 17-22.

Niezrecki, C., R. Phillips, and M. Meyer. 2003. Acoustic detection of manatee vocalizations. *J. Acoust. Soc. Amer.* 114: 1640-1647.

Oppenheim, A.V., and R.W. Schafer. 1999. *Discrete-Time Signal Processing*, 2nd ed. Prentice-Hall, Englewood Cliffs.

Stafford, K.M., C.G. Fox, and D.S. Clark. 1998. Long-range acoustic detection and localization of blue whale calls in the northeast Pacific Ocean. *J. Acoust. Soc. Am.* 104: 3616-3625.

Stafford, K.M., S.L. Nieu Kirk, and C.G. Fox. 2001. Geographic and seasonal variation of blue whale calls in the North Pacific. *J. Cetacean Res. Manage.* 3: 65-76.

Thompson, P.O. and W.A. Friedl. 1982. A long term study of low frequency sounds from several species of whales off Oahu, Hawaii.

Cetology 45: 1-19.

Vanderlaan, A.S.M., A.E. Hay, and C.T. Taggart. 2003. Characterization of North Atlantic right-whale (*Eubalaena glacialis*) sounds in the Bay of Fundy. *IEEE J. Oceanic Engr.* 28: 164-173.

Van Trees, H.L. 1968. *Detection, Estimation, and Modulation Theory, Vol. I.* Wiley: New York.

Waite, J.M., K. Wynne, and D.K. Mellinger. 2003. Documented sighting of a North Pacific right whale in the Gulf of Alaska and post-sighting acoustic monitoring. *Northwestern Naturalist* 84: 38-43.

Wiggins, S. 2003. Autonomous Acoustic Recording Packages (ARPs) for long-term monitoring of whale sounds. *Mar. Tech. Sci. J.* 37: 13-22.

Wiggins, S.M., M.A. McDonald, L.M. Munger, J.A. Hildebrand, and S.E. Moore. 2004. Waveguide propagation allows range estimates for North Pacific right whales in the Bering Sea. *Canad. Acoust.* 32: 146-154.

PYROK ACOUSTEMENT ACOUSTICAL PLASTERS CEILING AND WALL FINISHES

Designers and owners choose Pyrok Acoustement for their school and university projects because Pyrok provides:

- Decorative plaster finishes
- Superior sound absorbing performance
- Resistance to damage
- Low life cycle cost
- Non-combustible formulation



Perrot Memorial Library, Old Greenwich, CT

CONTACT:

Howard Podolsky, Pyrok, Inc.
914-777-7070
E-mail: info@pyrokinc.com or
www.pyrokinc.com



William Hart High School, Newhall, CA



High School for Physical City,
New York, NY



Carmel Valley Recreation Center,
San Diego, CA