Appendix A. Automated Detection and Classification of Marine Mammal Vocalizations and Anthropogenic Noise

A.1. Introduction

This appendix describes the methods developed by JASCO Applied Sciences Ltd. for automated detection of bearded seal calls, beluga whistles, bowhead moans, bowhead songs, minke whale calls, and walrus grunts within the data collected during the winter 2013−2014 and summer 2014 seasons of the Acoustic Monitoring Program in the northeastern Chukchi Sea. The algorithms, developed by JASCO, and their performance are described.

Methods to automatically detect and classify marine mammal vocalizations[1] in digital acoustic recordings have been developed over several decades. The variability of the target vocalizations influences the performance of detection algorithms. Some species, such as fin and blue whales, produce highly stereotyped vocalizations, which are easier to detect automatically than are sounds that vary more. For these stereotyped vocalizations, template-matching methods such as matched filter (Stafford 1995) and correlation of spectrograms (Mellinger and Clark 1997, 2000, Mouy et al. 2009) are generally effective (Mellinger et al. 2007). Other species produce highly varied and complex tonal sounds that are more difficult to detect and classify. Such vocalizations generally require band-limited energy summation for detection, followed by statistical classification techniques to identify species (Fristrup and Watkins 1993, Oswald et al. 2003). Several classification methods have been investigated for belugas (Clemins and Johnson 2006, Mouy et al. 2008), dolphins (Oswald et al. 2007), humpback whales (Abbot et al. 2010), elephants (Clemins et al. 2005), and birds (Kogan and Margoliash 1998).

Acoustical surroundings also influence how well detection algorithms perform. Increased ambient noise reduces the signal-to-noise ratio of vocalizations, making them harder to detect and classify. Noise generated by anthropogenic activities such a shipping and seismic exploration, or by weather such as wind, rain, and waves, could be mistaken as biological. The sound propagation characteristics of the study area can alter the spectral and temporal structure of received vocalizations, which can interfere with detection and classification algorithms that have worked well in a different propagation environment. The presence of other marine animals vocalizing in the frequency band of interest also greatly increases the risk of misclassification. How influential these factors are can vary with time. Consequently, methods shown to be successful for a specific location, season, and species might not be successful under different circumstances.

The program recordings contain vocalizations produced by several species of marine mammals, including bowhead (*Balaena mysticetus*), beluga (*Delphinapterus leucas*), gray (*Eschrichtius robustus*), fin (*Balaenoptera physalus*), and killer (*Orcinus orca*) whales, walrus (*Odobenus*

[1] Although many sounds made by marine mammals do not originate from vocal cords, the term “vocalization” is used as a generic term to cover all sounds produced by marine mammals that are discussed in this report. The term “call” is used synonymously for brevity.

rosmarus), and various ice seals. Several of these species vocalize in the same frequency bands, and their vocalizations can occur around the same time annually. For instance, bowheads and walrus vocalizations can be of similar durations and frequency ranges. While an experienced human analyst can usually distinguish between those vocalizations, it is not simple to create a computer algorithm to do the same.

Multiple sources contribute to ambient noise in the eastern Chukchi Sea. In winter, ice noise is highly problematic for automated detection algorithms—ice-cracking sounds can be emitted at surprisingly regular intervals, which resemble walrus knocks. Ice squeaking sounds are often in the frequency range of beluga vocalizations. Detection algorithms, therefore, must be well adapted to the variable and overlapping vocalizations of the species that frequent the northeastern Chukchi Sea as well as robust against the surrounding background noise. Because many terabytes of data were collected during the Acoustic Monitoring Program, the automated analysis methods must also be computationally efficient, with computing times taking no less than five times real time (per processor).

A.2. Bowhead and Beluga Call Detection and Classification

The bowhead acoustic repertoire includes low-frequency moans (< 1 kHz) produced in summer and higher frequency, more complex songs produced in fall and early winter (Delarue et al. 2009). Belugas produce tonal whistles in the 1–8 kHz frequency band (Karlsen et al. 2002).

Because these three sound types are produced in different frequency bands, three unique detectors and classifiers were created for:

- Bowhead winter and fall songs
- Bowhead summer moans
- Beluga whistles

To optimize performance on the call type of interest, each detector has unique spectrogram settings. The output of each detector was run through its associated classifier.

The detection/classification process consists of the following steps:

1. Creating the normalized spectrogram.
2. Extracting the time-frequency contours using the tonal detector developed by Mellinger et al. (2011).
3. Extracting 46 features from each contour to create binary random forest models.
4. Classifying the contours as either “target species” (bowhead or beluga) or “other” with the random forest models.
5. Post-processing of bowhead moans and songs to combine parts of single calls that were detected separately.

Once random forest models were created for bowhead moans, bowhead songs, and beluga whistles, they were tested on the test datasets described in Section A.6. The detection/classification process is described in detail in the following sections.

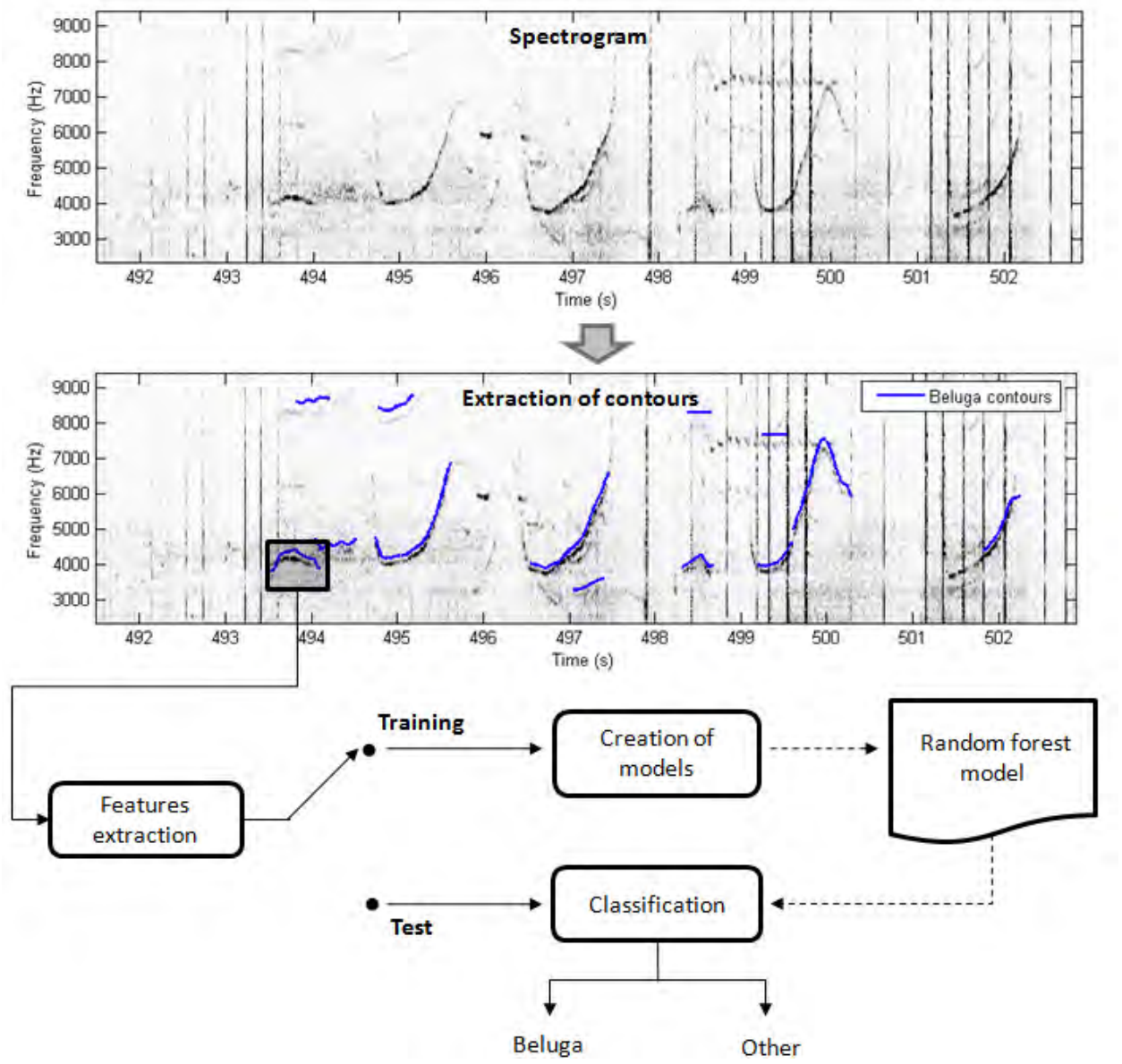

Figure A-1. Steps in the detection/classification process.

A.2.1. Step 1: Spectrogram Processing

The first step of the detection process was the calculation of the spectrogram. Spectrogram resolutions differed for each species to ensure accurate time-frequency representation of the calls (Table A-1). To attenuate long spectral rays in the spectrogram due to vessel noise and to enhance weaker transient biological sounds, the spectrogram was normalized in each frequency band (i.e., each row of the spectrogram) with a split-window normalizer. The size of the window and the notch of the normalizer are indicated in Table A-1. For the processing of beluga whistles, the spectrogram was smoothed by convolving it with a 2-D Gaussian kernel (Gillespie 2004). Gaussian smoothing was not used for analyzing bowhead calls because it did not improve the performance of the contour extraction.

Table A-1. Spectrogram parameters for each call type.

	Bowhead winter songs	Bowhead summer moans	Beluga whistles
Analysis frame size (samples)	4096	4096	1024
Overlap between frames (samples)	3500	3500	896
FFT size (samples)	16,384	16,384	1024
Window function	Hanning	Hanning	Blackman
Normalizer window size (s)	1.5	1.5	0.7
Normalizer notch size (s)	0.4	0.4	0.1
Gaussian kernel size (bins)	n/a	n/a	3×3

A.2.2. Step 2: Contour Extraction

Vectors representing the time-evolution of the fundamental frequency of marine mammal calls (referred to as "contours") were extracted from the spectrograms with the MATLAB version of a tonal detector developed by Mellinger et al. (2011). This tonal detector is implemented in the latest version of the widely-used Ishmael acoustic analysis software (Mellinger 2001). The algorithm works as follows, based on user-defined parameters (chosen empirically, Table A-2):

1. Candidate frequency peaks were identified for each time slice of the spectrogram in the frequency band $[f_0, f_1]$. Peaks of height h (dB) above the noise threshold (defined as the percentile P_{bg} of the spectrum values) that are the highest point in their neighborhood (n Hz wide) were selected.
2. Successive peaks differing in frequency by less than f_d were connected.
3. To accurately follow simultaneous calls, the location of the next candidate peak was estimated by fitting a line to the most recent k seconds of the contour and looking for spectral peaks where the line continues.
4. Candidate contours must persist for a minimum duration d.

Figure A-1 shows an example of contours extracted from a recording containing beluga whistles.

Table A-2. Contour extraction parameters for each call type.

Symbol	Description	Bowhead winter songs	Bowhead summer moans	Beluga whistles
P_{bg}	Percentile for estimating background noise	50	50	50
h	Height above that estimate (dB)	2	2	1.2
n	Neighborhood width (Hz)	50	50	250
fd	Frequency difference from one step to the next (Hz)	25	25	300
d	Minimum duration (s)	0.5	0.5	0.3
k	Duration for estimating next spectral peak location (s)	0.2	0.2	0.2
f_0	Minimum frequency (Hz)	1000	50	50
f_1	Maximum frequency (Hz)	1000	50	8000

A.2.3. Step 3: Feature Extraction

Using custom MATLAB software, 46 features were measured from each extracted time-frequency contour. These features describe the frequency content, duration, and shape of the contour (slopes, number of inflection points, etc., Table A-3).

Table A-3. The 46 features measured from each time-frequency contour.

Feature	Definition
Beginning sweep	Slope at the beginning of the call (1 = positive, −1 = negative, 0 = flat)
Beginning up	Binary variable: 1 = beginning slope is positive, 0 = beginning slope is negative
Beginning down	Binary variable: 1 = beginning slope is negative, 0 = beginning slope is positive
End sweep	Slope at the end of the call (1 = positive, −1 = negative, 0 = flat)
End up	Binary variable: 1 = ending slope is positive, 0 = ending slope is negative
End down	Binary variable: 1 = ending slope is negative, 0 = ending slope is positive
Duration	Call duration (s)
Beginning frequency	Frequency at start of call (Hz)
End frequency	Frequency at end of call (Hz)
Minimum frequency, f_{min}	Minimum frequency (Hz)
Maximum frequency, f_{max}	Maximum frequency (Hz)
Frequency range	$f_{max}–f_{min}$ (Hz)
Mean frequency	Mean of frequency values (Hz)
Median frequency	Median of frequency values (Hz)
Standard deviation frequency	Standard deviation frequency values (Hz)
Frequency spread	Difference between the 75th and 25th percentiles of the frequency
Quarter frequency	Frequency at one-quarter of the duration (Hz)
Half frequency	Frequency at one-half of the duration (Hz)
Three-quarter frequency	Frequency at three-quarters of the duration (Hz)
Center frequency, f_c	$(f_{max}–f_{min})/2 + f_{min}$
Relative bandwidth	$(f_{max}–f_{min})/f_c$
Maxmin	f_{max}/f_{min}
Begend	Beginning frequency/end frequency
Steps	Number of steps (≥ 10% increase or decrease in frequency over two contour pts)
Inflection points	Number of inflection points (changes from positive to negative slope or *vice versa*)
Max delta	Maximum time between inflection points
Min delta	Minimum time between inflection points
Maxmin delta	Max delta/Min delta
Mean delta	Mean time between inflection points
Standard deviation delta	Standard deviation of the time between inflection points
Median delta	Median of the time between inflection points
Mean slope	Overall mean slope
Mean positive	Mean positive slope
Mean negative	Mean negative slope
Mean absolute	Mean absolute value of the slope
Ratio posneg	Mean positive slope/Mean negative slope
Percent up	Percentage of the call having positive slope
Percent down	Percentage of the call having negative slope
Percent flat	Percentage of the call having zero slope
Up-down	Number of inflection points going from positive to negative slope
Up-flat	Number of times the slope changes from positive to zero
Flat-down	Number of times the slope changes from zero to negative
Step-up	Number of steps with increasing frequency
Step-down	Number of steps with decreasing frequency
Step-duration	Number of steps/Duration
Inflection-duration	Number of inflection points/Duration

A.2.4. Step 4: Classification

A random forest classifier was created for each call type (bowhead winter songs, bowhead summer moans, and beluga whistles). A random forest is a collection of decision trees that grow using binary partitioning of the data based on the value of one of the 46 features (see Table A-3) at each branch, or node. Randomness is injected into the tree-growing process by choosing the feature to use as the splitter based on a random subsample of the features at each node (Breiman 2001). Each of these random forests was a binary classifier, so contours were classified as "target species" (i.e., bowhead or beluga whale) or "other".

The number of decision trees to include in each random forest was determined by empirical trials on datasets of calls extracted from annotated recordings. Recordings from prior years were used to train and optimize the random forests: winter 2008–2009 program data for the bowhead winter song and beluga whistle detectors, and summer 2009 program data for the bowhead summer moan detector. Contours were detected and extracted based on parameters specific to bowhead or beluga sounds (Table A-2).

Sample sizes for each trial dataset are given in Table A-4. First, these datasets were randomly sampled so each class ("target species" and "other") had equal sample sizes. Sampling was performed so that the proportion of species and call-types within species in the "other" class reflected those in the full dataset. Next, a random forest analysis was run on the sampled data. The sampling and random forest analyses were each repeated 100 times. The output for each random forest analysis included out-of-bag (OOB) error estimates for forests of 1–800 trees. To calculate OOB errors, each tree was grown using about two-thirds of the trial data. The remaining third of the trial data was used as the OOB test data, which was used to evaluate the performance of the tree. The OOB error estimates were averaged over 100 runs (Figure A-2). The number of decision trees to include in the random forest was when the OOB error approached its asymptote, because after this point adding more trees did not result in significantly better classifications. Based on these analyses, all three random forests had 300 decision trees.

Table A-4. Sample size of the trial datasets used to train and optimize the random forest classifiers for each call type.

Class	Winter 2008–2009 Beluga whistles	Winter 2008–2009 Bowhead songs	Summer 2009 Bowhead moans
Beluga	1295	24	0
Bowhead	2837	3989	754
Bearded seal	20,331	17,887	269
Non-biological noise	9443	6491	536
Ribbon seal	530	0	0
Unknown	864	1148	1177
Walrus	483	199	625
Killer whale	0	0	13

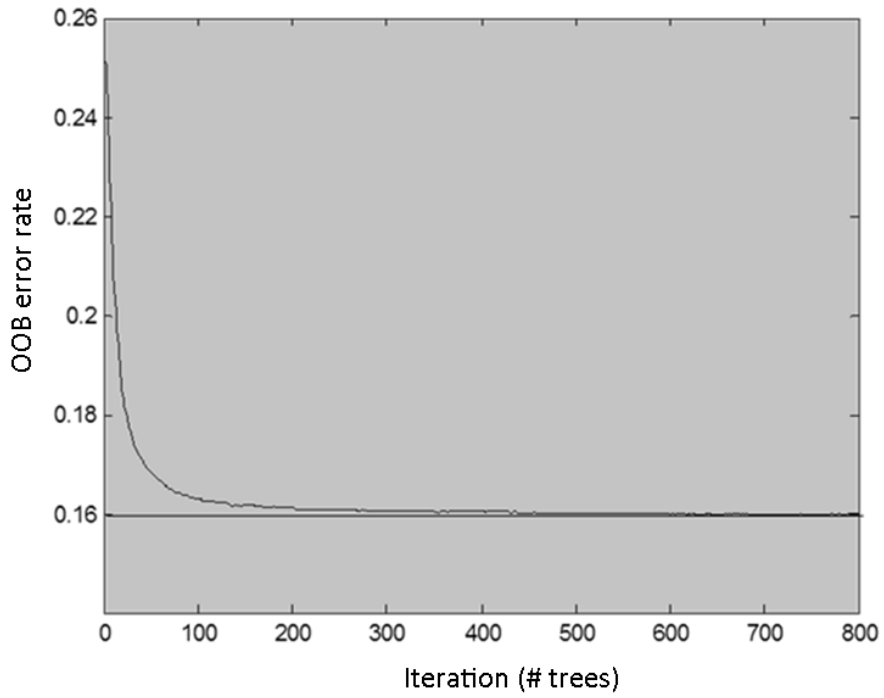

Figure A-2. Out-of-bag (OOB) error rates averaged over 100 random forest runs (example of the beluga whistle classifier).

Another output of the random forest analysis is the Gini importance index (Breiman et al. 1984), which measures how strongly each feature contributes to the random forest model predictions. The optimal subset of features included in each random forest was based on this importance index. Feature importance was averaged over all 100 runs, which were described above (Figure A-3). The three random forests included the features most important to the model predictions (Table A-5).

Table A-5. Features included in bowhead moan, bowhead song, and beluga whistle random forests, listed in order of importance to the model.

Bowhead moan	Bowhead song	Beluga whistle
Minimum frequency	Maximum frequency	Mean frequency
Median frequency	Center frequency	End frequency
Mean frequency	Beginning frequency	Median frequency
Three-quarter frequency	Mean frequency	Three-quarter frequency
End frequency	End frequency	Center frequency
Half frequency	Mean slope	Half frequency
Quarter frequency	Median frequency	Maximum frequency
Beginning frequency	Quarter frequency	Quarter frequency
Duration	Three-quarter frequency	Minimum frequency
Center frequency	Half frequency	Beginning frequency
Mean negative slope	Mean absolute slope	

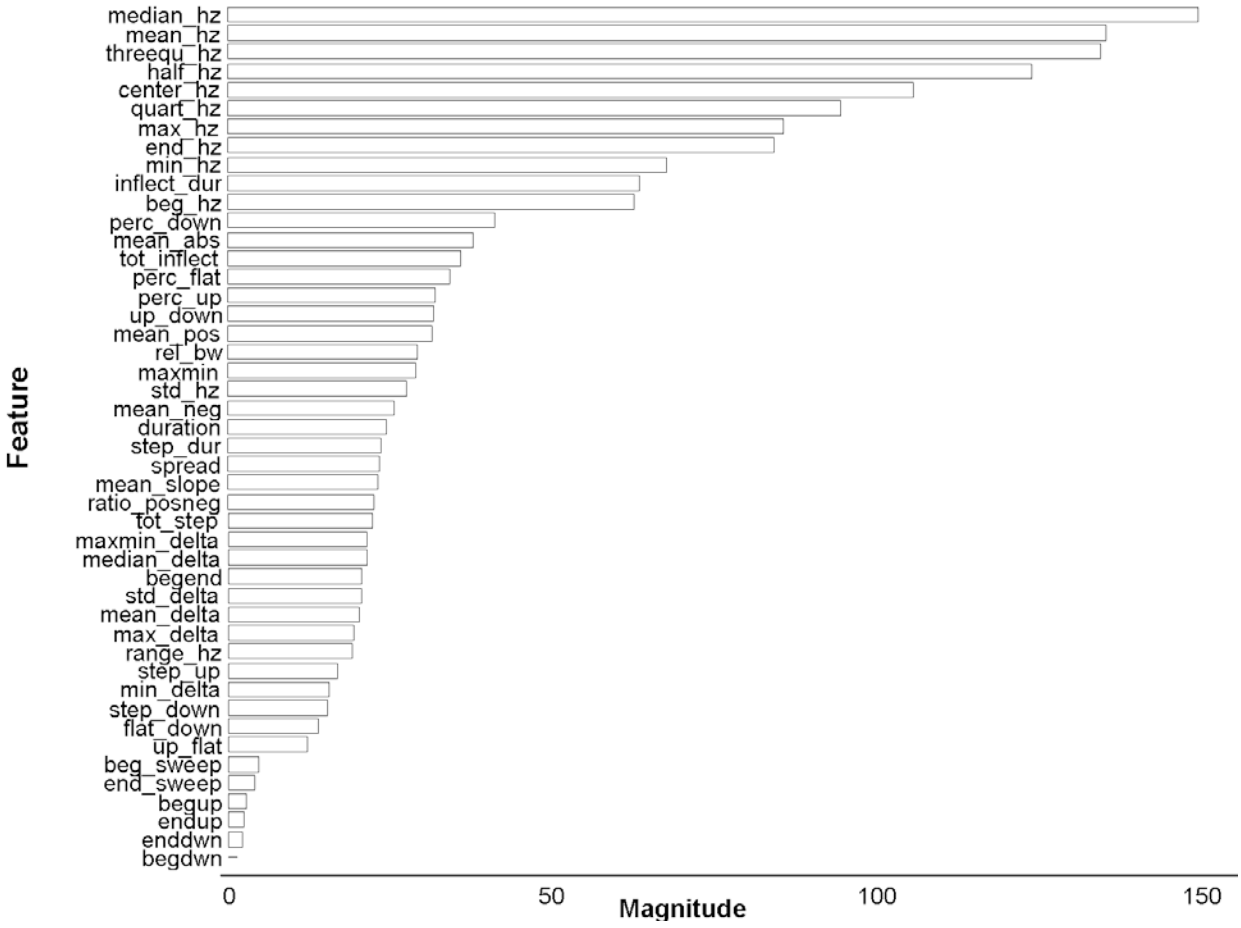

Figure A-3. Gini importance indices; averaged over 100 random forest runs.

A.2.5. Step 5: Post-Processing

Bowhead calls recorded in the winter program generally consisted of several harmonics, which the automated detector considered separate calls, thus overestimating the number of calls in the recordings. To avoid this, all bowhead detections that overlapped in time were merged to form a single detection. Only detections occurring below 300 Hz were considered. No post-processing was performed on beluga detections.

A.3. Walrus Grunt Detection and Classification

The algorithm first calculated the spectrogram and normalized it for each frequency band. Then the spectrogram was segmented between 10 and 1000 Hz to define acoustic events in the spectrogram. For each event, a set of features representing salient characteristics of the spectrogram were extracted. Extracted features were presented to a five-class random forest classifier to determine the class of the sound detected (i.e., “walrus grunt”, “seismic”, “bowhead”, “bearded seal”, or “other”). During the training phase, features of known sounds (i.e., manual annotations) were extracted to create the random forest model. Figure A-4 illustrates the detection process.

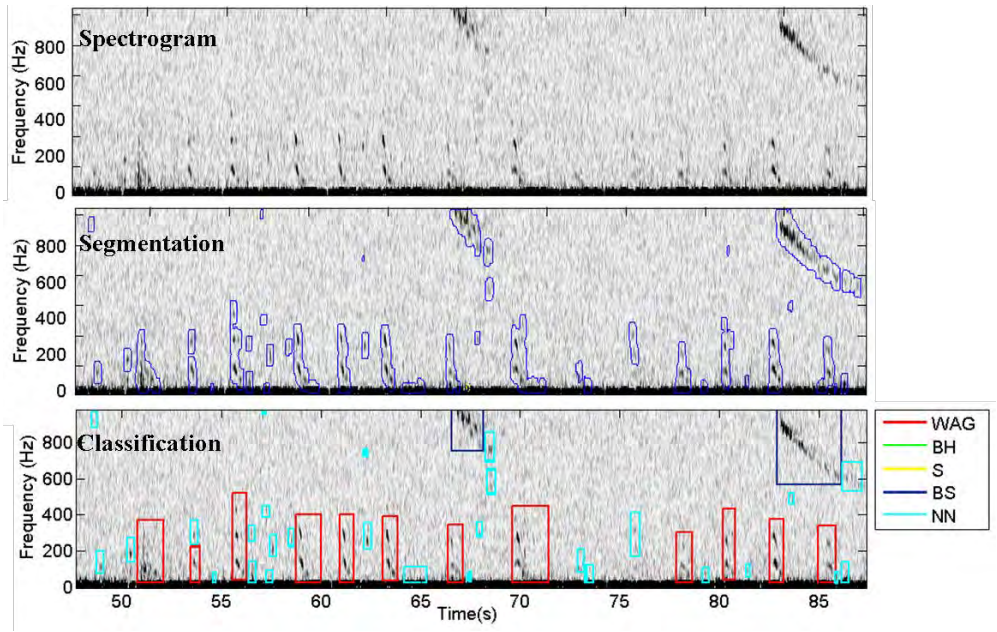

Figure A-4. Steps of the walrus grunt detector. Blue lines in the middle panel indicate edges of detected objects. Colors in the bottom panel indicate whether objects were classified as walrus (WAG, red), bowhead (BH, green), seismic (S, yellow), beaded seal (BS, dark blue), or noise (NN, light blue).

A.3.1. Step 1: Spectrogram Processing

The spectrogram resolution was chosen to ensure accurate time-frequency representation of the walrus grunts (Table A-6). The spectrogram was normalized by a split window normalizer using a window of 3 s and a notch of 1 s (Struzinski and Lowe 1984).

Table A-6. Spectrogram parameters used in the walrus grunt detector.

Spectrogram parameters	
Analysis frame size (samples)	1024
Overlap between frames (samples)	700
FFT size (sample)	2048
Window function	Blackman

A.3.2. Step 2: Spectrogram segmentation

The spectrogram is segmented by calculating the local variance of energy values on a 2-dimensional kernel. The local variance was calculated twice using different sizes of kernel. The first pass was performed using a kernel of 0.1 s by 50 Hz. Areas of the spectrogram with a local variance less than 0.6 were set to zero (variance units: energy^2). The second pass used a kernel of size 0.3 s by 100 Hz. Areas of the spectrogram with a local variance less than 0.6 were set to zero (variance units: energy^2). The first pass defined each component of the transients at small scale, while the second pass removed small noise objects and group together calls with harmonics. Edges of the remaining area of the spectrogram were defined using the Moore Neighborhood algorithm (Ainslie and McColm 1998). Figure A-4 (middle panel) shows an

example of the segmentation process. Finally, only objects longer than 100 ms and larger than 20 Hz were kept for classification.

A.3.3. Step 3: Feature Extraction

Features for each object were extracted on a time-frequency box that contains 95% of the energy of the initial object (red box in Figure A-5; top panel). Each object was represented by 40 features, several of which were calculated following Fristrup and Watkins (1993) and Mellinger and Bradbury (2007), using the spectrogram, frequency envelope, and amplitude envelope of the signal (Figure A-5). The frequency envelope is the sum of the spectrogram amplitudes for each frequency. The maximum of the frequency envelope was normalized to 1. The amplitude envelope is the sum of the spectrogram amplitude values for each time step. The frequency and amplitude envelopes were interpolated to have a resolution of 1 Hz and 1 ms respectively. Features include the following:

- Median frequency, f_{med}: Based on the frequency envelope. The cumulative sum of the spectrum was calculated by moving from low to high frequencies. The median frequency is the frequency at which the cumulative energy reaches 50% of the total energy (green dashed line in Figure A-5).
- Spectral inter-quartile range: Calculated by defining the 25th percentile of the energy on each side of the median frequency (dashed blue lines in Figure A-5). Each quartile was defined as frequency for which the cumulative energy calculated from the median frequency equaled 25% of the total energy. The spectral inter-quartile range is the difference between the higher (f_{Q3}) and lower quartiles (f_{Q1}).
- Spectral asymmetry: Skewness of the spectral envelope calculated as $(f_{Q1} + f_{Q3} – 2f_{med})/(f_{Q1} + f_{Q3})$.
- Spectral concentration : Calculated by ranking amplitude values of the spectral envelope from largest to smallest. The cumulative sum of ranked amplitude values was computed beginning with larger values until 50% of the total energy was reached. The lowest frequency index included in the additive set was considered the minimum; the highest index was the maximum, with their difference providing the spectral concentration (red box in Figure A-5).
- Maximum frequency peak: Frequency of the highest amplitude peak in the spectral envelope (red dot in Figure A-5).
- Maximum frequency peak width: Width (Hz) of the maximum frequency peak measured at the point where amplitude values on each side of the peak reached the 75th percentile of all the spectral envelope amplitude values (red vertical line in Figure A-5)
- Difference in Hz between the maximum frequency peak and the median frequency of the frequency envelope.
- Maximum and minimum frequency of the object.
- Frequency bandwidth and duration of the object.
- Variance and kurtosis of frequency envelope: These describe the distribution of the amplitude in the spectral envelope (Balanda and MacGillivray 1988).

- Frequency modulation index: Calculated as follows:
  - First, the maximum frequency of the maximum amplitude peak was extracted for each time slice of the spectrogram. Frequency values of the selected peaks were stored in the vector F_{max}, and their associated energy values in the vector E_{max}. Only peaks with an amplitude value exceeding the median amplitude of the spectrogram were considered (white dots in Figure A-5a).
  - Second, the weighted maximum frequency offset vector O was defined as $O = (F_{max} – X_{med}) \cdot E_{max} / max(E_{max})$, where X_{med} is a scalar representing the median frequency of the vector F_{max}. The frequency modulation index was defined as the standard deviation of the vector O.
- Frequency and correlation value of the maximum peak in the autocorrelation function calculated on the frequency envelope.
- Temporal inter-quartile range: same as the frequency inter-quartile range but calculated on the time envelope.
- Temporal asymmetry: same as the frequency asymmetry but calculated on the time envelope.
- Temporal concentration: same as the frequency envelope but calculated on the time envelope.
- Variance and kurtosis of time envelope.
- Period and correlation value of the maximum peak in the autocorrelation function calculated on the time envelope.
- Several features based on the contour representing the evolution in time of the median frequency. These include the number of inflection points, upsweep and downsweep rates, standard deviation, and variance of frequency values.

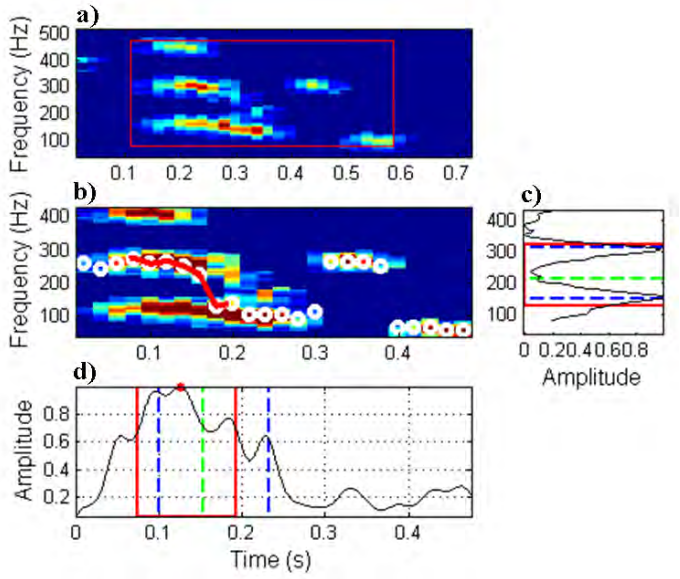

Figure A-5. Extraction of features used in the walrus grunt classifier: (a) Spectrogram of the original object detected. The red rectangle indicates the zone including 95% of the energy; (b) spectrogram of the resized object including 95% of the energy. White dots indicate the median at each time step and the red line indicates the median contour; (c) frequency envelope (black line), with the median frequency (green line), the upper and lower quartiles (blue lines), and the spectral concentration (red box); (d) amplitude envelope with the median (green line), the upper and lower quartiles (blue lines), and the temporal concentration (red box).

A.3.4. Step 4: Classification

Classification was performed using a random forest classifier (Breiman 2001), which was trained using all manual annotations in recordings from the summer 2011 Acoustic Monitoring Program. The random forest was defined with these four classes: “walrus grunt”, “seismic”, “bowhead”, “bearded seal”, and “other”. Training the classifier, optimizing the number of decision trees in the forest, and selecting the most relevant features based on the Gini index, were performed using the same process described for bowhead and beluga whale call detection (Section A.2).

A.4. Bearded Seal Call Detection

The automated detection and classification of bearded seal calls is performed in four steps:

1. Calculation and binarization of the spectrogram
2. Definition of time-frequency objects
3. Extraction of features
4. Classification

A.4.1. Step 1: Spectrogram Processing

The first step of the detection process was calculating the spectrogram. Table A-7 lists the spectrogram parameters. To attenuate long spectral rays in the spectrogram due to vessel noise,

and to enhance weaker transient biological sounds, the spectrogram was normalized in each frequency band (i.e., each row of the spectrogram) with a median normalizer. Table A-7 lists the size of the window used by the normalizer. The normalized spectrogram was binarized by setting all the time-frequency bins that exceed a normalized amplitude of 4 (no unit) to 1 and the other bins to 0.

Table A-7. Spectrogram parameters used in the bearded seal call detector.

Spectrogram parameters	Bearded seal calls
Analysis frame size (samples)	4096
Overlap between frames (samples)	3072
FFT size (samples)	4096
Window function	Reisz
Normalizer window size (s)	120
Binarization threshold (no units)	4

A.4.2. Step 2: Definition of Time-Frequency Objects

The second step of the detection process consisted of defining time-frequency objects (or events) by associating contiguous bins in the binary spectrogram. The algorithm implemented is a variation of the flood-fill algorithm (Nosal 2008). Every spectrogram bin that equals 1 and is separated by less than three bins in both time and frequency are connected. Figure A-6 illustrates the search area used to connect spectrogram bins. The bin connection process moves from oldest data to newest and from lowest frequency to highest. Each group of connected bins is referred to as a time-frequency object. A spectrogram bin can only belong to one time-frequency object.

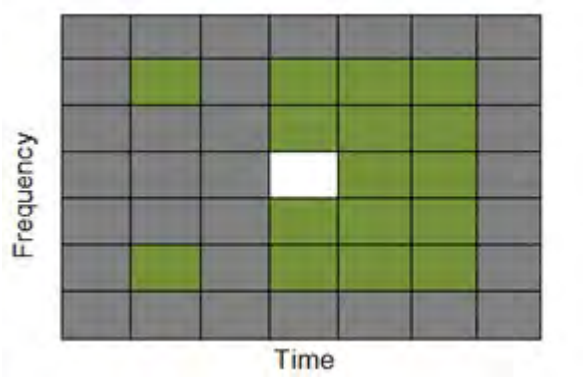

Figure A-6. Illustration of the search area used to connect spectrogram bins. The white square represents a bin of the binary spectrogram equaling 1. The green squares represent the bins to which it could potentially be connected. The algorithm advances from left to right so gray cells left of the test cell need not be checked; however, checking the far left cells could join broken contours.

Because time-frequency objects are sensitive to noise generated by small pleasure craft or fishing vessels near recorders—they can generate many time-frequency objects that might be mistaken for marine life calls—to reduce false detections a vessel detector is incorporated into the time-frequency event definition process. Vessel noise is considered detected when at least five

frequencies have detected contours for 5 s. Files with at least two vessel detections are not processed further.

A.4.3. Step 3: Feature extraction

The third step consists of representing each of the time-frequency objects extracted in the previous step by a set of features. Features of the time-frequency objects were defined by:

- Start time (date)
- Duration (s)
- Minimum frequency (Hz)
- Maximum frequency (Hz)
- Bandwidth (Hz)

A.4.4. Step 4: Classification

The final step consisted of classifying the time-frequency objects by comparing their features against a call definition dictionary that defines the features of the vocalizations present in the Chukchi Sea based on the literature and on analysts' observations. In the present study, only bearded seal calls were represented in the dictionary (Table A-8). The classification process can handle vocalizations made of several time-frequency objects, such as vocalizations with harmonics ("Multi-Frequency-Components") and vocalizations made of a succession of time-frequency objects such as seal trills and groups of beluga, dolphin, or beaked whale whistles ("Multi-Time-Components").

Vocalizations in the dictionary are defined by the following features:

- Minimum frequency
- Maximum frequency
- Minimum duration: at least one spectrogram time slice.
- Maximum duration
- Minimum bandwidth
- Maximum bandwidth
- Multi-Frequency-Component (Boolean): for call types where contours should be grouped in frequency, with some time overlap before applying the frequency, duration, and bandwidth constraints. Each contour that is added to the multi-component contour has the following constraints applied:
- minComponentDuration: minimum duration for a contour to be added to the multi-component contour.
- minComponentBW: minimum bandwidth for a contour to be added to the multi-component contour.

- Minimum and maximum frequencies: as per the global definition.
- Multi-Time-Component (Boolean): for call types where contours should be grouped in time before applying the frequency, duration, and bandwidth constraints. Each contour that is added to the multi-time-component contour has the following constraints applied:
- minTimeComponentDuration: minimum duration for a contour to be added to the multi-time-component contour.
- minTimeComponentBW: minimum bandwidth for a contour to be added to the multi-time-component contour.
- Minimum and maximum frequencies: as per the global definition.

Table A-8. Definitions for the time-frequency features of bearded seal calls in the Chukchi Sea in the summer and in the winter.

	Call Type	Min/Max frequency (Hz)	Min/Max duration (s)	Min/Max bandwidth (Hz)	Min/Max sweep rate	Multi-Frequency-component settings	Multi-time-component settings
Winter Calls	Full Trill	250/5000	5/60	500/–	−100/−10	Min BW=30 Max BW=200 Min Dur=0.5 Max Dur=5 MaxFreqShift=100	0
	Trill end	250/1200	10/60	100/–	−50/−5	Min BW=20 Max BW=100 Min Dur=0.5 Max Dur=8 MaxFreqShift=100	0
Summer Calls	Downsweep	200/1500	0.6/10	38/–	−200/−20	N/A	0
	Upsweep	200/1500	0.6/4.5	100/–	50/250	N/A	0

Figure A-7 shows a block diagram of the several stages of the classification algorithm. The algorithm consists of two loops. The outer loop iterates through all the time-frequency objects. For each time-frequency object that has not yet been classified, the object's features are compared to each call in the dictionary. If the call is a multi-frequency-component or multi-time-component type, the list of time-frequency objects is searched for unsorted objects that meet the multi-components settings (see Table A-8).

The total time-frequency object duration, minimum and maximum frequencies, and frequency bandwidth are compared to the call's definitions in the dictionary. If the object's features fall within the call type's bounds, then the bandwidth (BW_i) and duration (T_i) indices are computed:

$$BW_i = \frac{BW_{object}}{BW_{dictionary}} \quad T_i = \frac{T_{object}}{T_{dictionary}}$$

If either of these indices exceeds the empirically chosen threshold of 1.5 times the current best index, then the current best-match call type is updated. The 1.5 threshold for updating the best-match call type means the algorithm prefers call types that are defined earlier. Therefore, if, for a particular recording, killer whales are more likely to occur than humpbacks, the killer whale call definitions should occur first in the mammalContours.xml definition file.

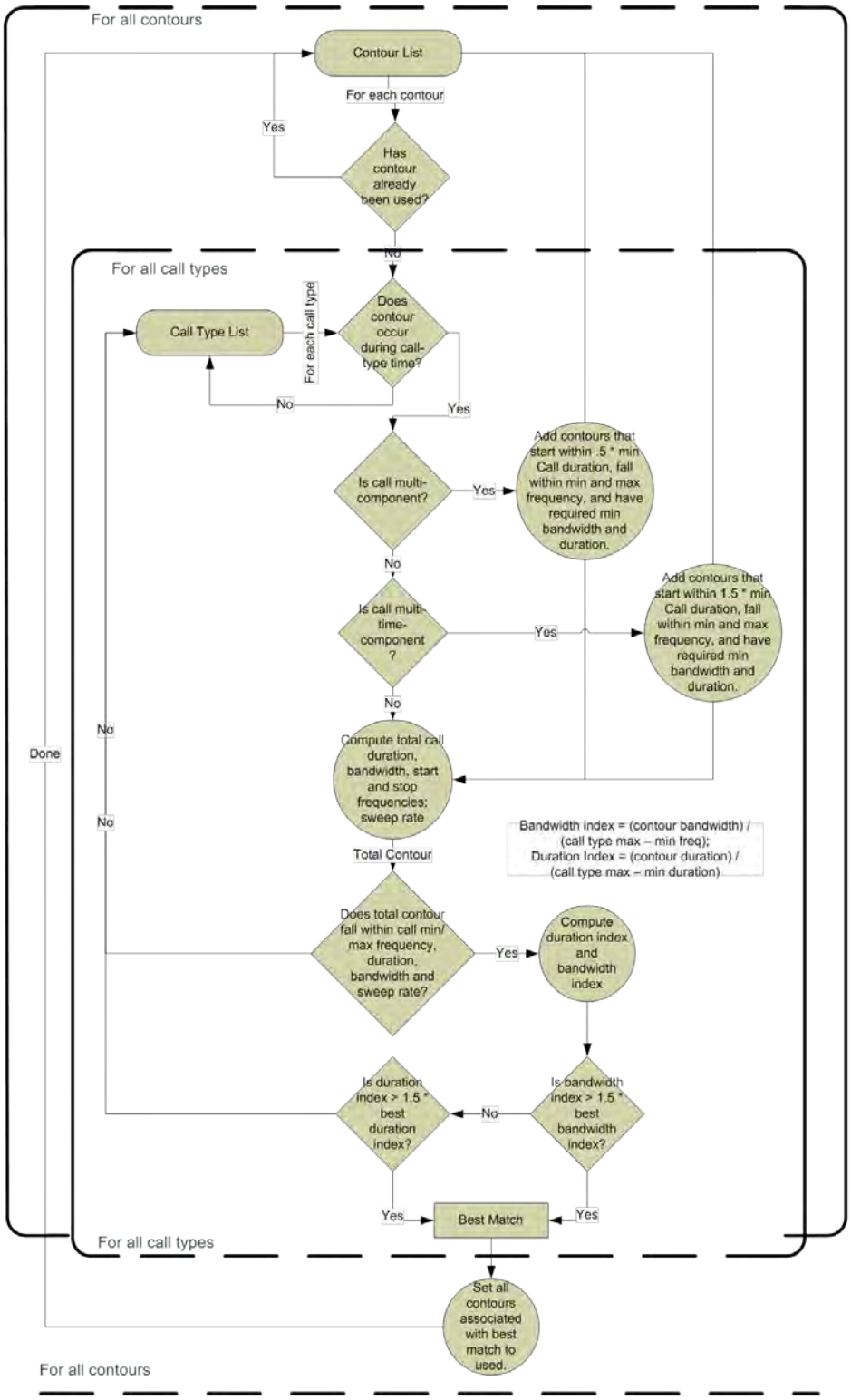

Figure A-7. Block diagram of the classification algorithm.

The classification algorithm also implements a time-based filter. Because the classification algorithm is intended to count calls of species expected to be in an area, it is reasonable to limit

the algorithm to those species. For instance, bowhead calls won't be detected before 1 Sep or after 1 Jan in the Chukchi Sea. Manual analysis is used to detect extra-limital species and unusual detections as a function of time. Figure A-8 shows an example of detection and classification of bearded seal calls.

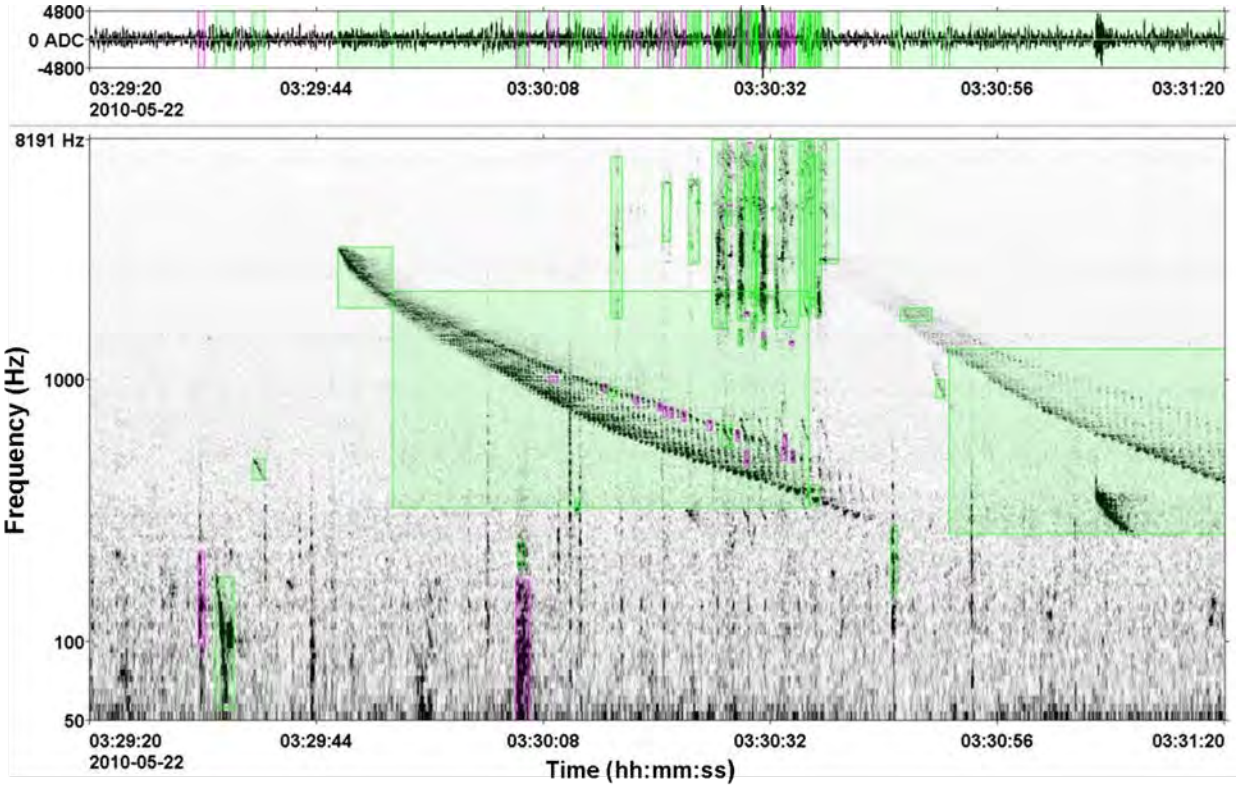

Figure A-8. (Top) Pressure in digital units and (bottom) spectrogram of bearded seal trills (500–200 Hz; downsweeps in center) detected using the multi-time-component contour type. Beluga and bowhead calls are also visible in this figure (16 kHz sample rate, 4096 pt STFT, 1024 pt advance).

A.5. Minke Whale Detection

The spectrogram correlation technique used for detecting minke whale calls is based on Mouy et al. (2009) and was configured to detect the boing calls (Rankin and Barlow 2005):

The spectrogram is computed and normalized using a split-window normalizer with a 30 s window and a 2 s notch (Struzinski and Lowe 1984).

The spectrogram is then binarized by calculating the variance of energy values around each spectrogram bin on a time-frequency kernel of size 1 s by 10 Hz. Bins of the spectrogram with a local variance less than 0.4 and a normalized energy value less than 2 are set to zero. Remaining spectrogram bins are set to 1.

A set of synthetic binary time-frequency templates representing typical minke whale boing calls were created as successions of linear time-frequency segments defined by their starting frequency, ending frequency, duration, frequency width, frequency span, and silence duration before and after the call. Here two time-frequency template were used (Table A-9)

Table A-9. Parameters of the time-frequency templates used to detect minke whale boing calls.

	Template 1		Template 2			
Segment number	1	2	1	2	3	4
Segment start time (s)	1	1.8	1	1.8	1	1.8
Segment stop time (s)	1.8	3.5	1.8	3.5	1.8	3.5
Segment frequency width (Hz)	15	15	15	15	15	15
Segment start frequency (Hz)	1365	1390	1365	1390	1497	1517
Segment stop frequency (Hz)	1390	1390	1390	1390	1517	1517
Min. frequency of the template (Hz)	1300	1300	1300	1300	1300	1300
Max. frequency of the template (Hz)	1470	1470	1570	1570	1570	1570
Duration of the template	4.5	4.5	4.5	4.5	4.5	4.5

To create a detection function, a correlation index that measured how well the synthetic templates matched the binary spectrogram was defined for each time step of the spectrogram and for each of the templates. The occurrences of minke whale call detections were defined by parts of the detection function that exceeded an empirically chosen threshold, T_{detec}. Here $T_{detec} = 0.1$ for both templates.

Figure A-9 illustrates the detection process.

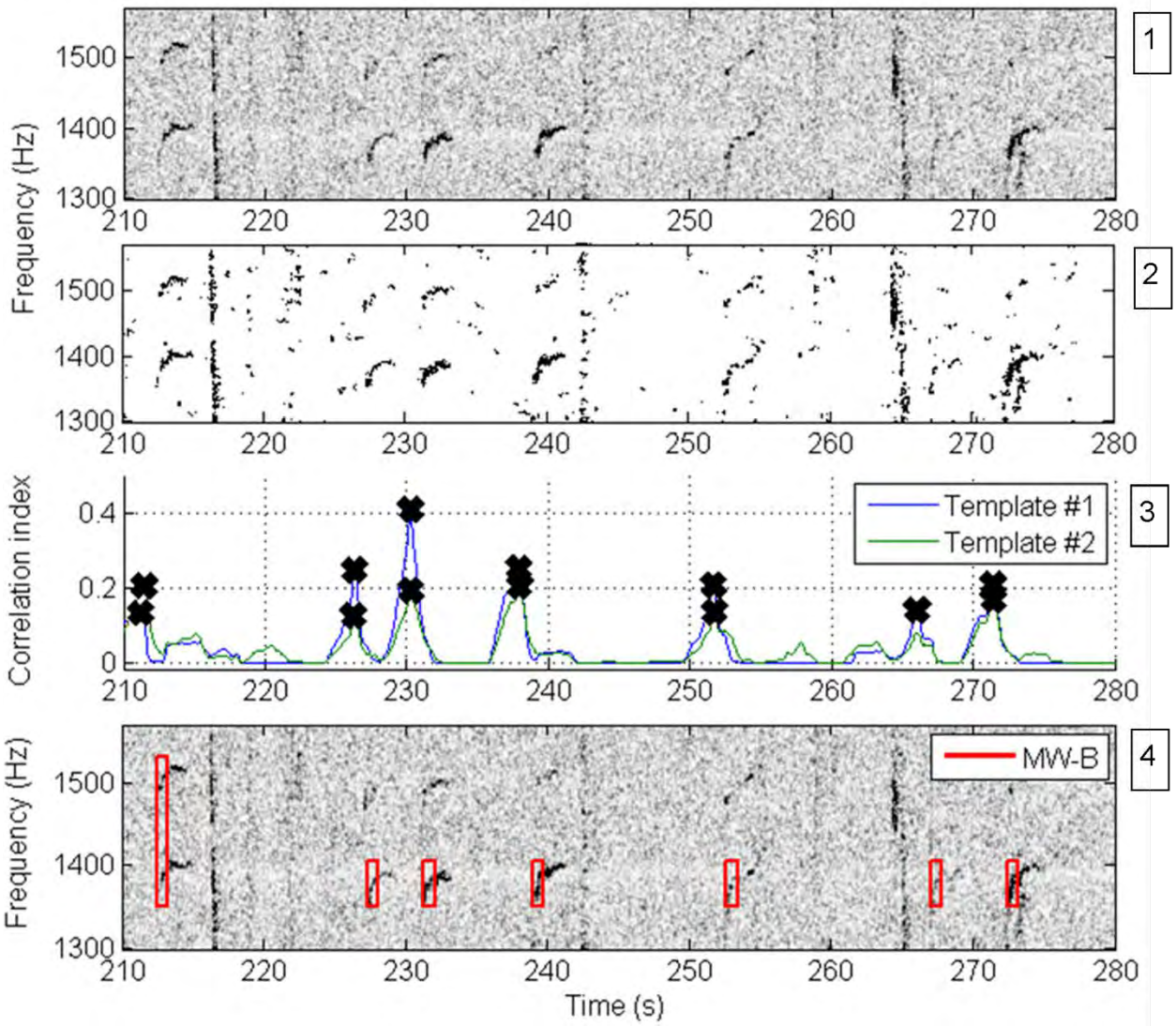

Figure A-9. Minke whale boing call detection process. Panel 1: Spectrogram representing 7 boing calls. Panel 2: Binarized spectrogram. Panel 3: Detection function for each template (black crosses indicated peaks above the detection threshold T_{detec}. Panel 4: Results of the detection process (red boxes indicate boing call detections).

A.6. Performance Evaluation

A.6.1. Test Datasets

The automated detectors/classifiers must be verified with a test dataset that represents the spatio-temporal variations of the marine mammal calls and background noise in the entire dataset. Since the acoustic environment in the eastern Chukchi Sea differs between winter and summer, a unique test dataset was used to test the detection/classification algorithms for each season. For the winter 2013–2014 Acoustic Monitoring Program data, marine mammal calls were fully manually-annotated in 186 samples adding up to 5.2 hours of recordings chosen from Stations B5, CL5, PL10, PL50, W10, W50, and WN40. For the summer 2014 Acoustic Monitoring

Program data, marine mammal calls were fully manually-annotated in 138 samples adding up to 10.6 hours of recordings chosen from Stations BGB, BGC, BGE, BGF, CLN120B, CLN40, CLN90B, KL01, PL30, PL50, PLN20, PLN40, PLN60, PLN80, W10, W30, W50, WN20, WN40, and WN80.

A.6.2. Performance Metrics

The decisions made by detectors/classifiers can be represented as a confusion matrix. The confusion matrix consists of four categories: true positives (*TP*), false positives (*FP*), true negatives (*TN*), and false negatives (*FN*). Table A-10 depicts the confusion matrix, where *E* is the signal event we want to detect/classify and $\overline{E}$ is a non-event that we want to ignore (i.e., noise). The definition of $\overline{E}$ varies depending on the detector/classifier.

Table A-10. Confusion matrix.

		True Result: E	True Result: $\overline{E}$
Detection/ classification result	E	TP	FP
	$\overline{E}$	FN	TN

A true positive (*TP*) corresponds to a signal of interest being correctly classified as such. A false negative (*FN*) is a signal of interest being classified as noise (i.e., missed). A false positive (*FP*) is a noise classified as a signal of interest (i.e., a false alarm). A true negative (*TN*) is a noise correctly classified as such.

The numbers of *TP*s, *FP*s, and *FN*s were calculated for each detector and test dataset by comparing the manual annotations of marine mammal calls (considered true results, i.e., ground truth) with the automated detections/classifications. Numbers of *FP*s, *TP*s, and *FN*s were calculated on all dataset samples containing annotations of the target call type. If a manually-annotated call was automatically detected/classified, then the detection was considered a *TP*, if undetected, it was a *FN*. Each automated detection occurring in the sample that did not correspond to a manually-annotated call was considered a *FP*.

A.6.3. Precision and Recall

To assess the performance of the detectors/classifiers, precision (*P*) and recall (*R*) metrics were calculated based on the numbers (*N*) of *TP*s, *FP*s, and *FN*s:

$$P = \frac{N_{TP}}{N_{TP} + N_{FP}} \quad R = \frac{N_{TP}}{N_{TP} + N_{FN}} \tag{1}$$

P measures exactness, and *R* completeness. For instance, a *P* of 0.9 for beluga means that 90% of the detections classified as beluga were in fact beluga calls, but says nothing about whether all

beluga vocalizations in the dataset were identified. An *R* of 0.8 for beluga means that 80% of all beluga calls in the dataset were classified, but says nothing about how many beluga classifications were wrong. Thus, a perfect detector/classifier would have $P = R = 1$. Neither *P* nor *R* alone can describe the performance of a detector/classifier on a given dataset; both metrics are required.

The advantages of the *P*-*R* metric over the True-Positive Rate (TPR) and False-Positive Rate (FPR) generally used in Receiver Operating Characteristic (ROC) curves include:

- The *P*-*R* metric is more adapted to skewed datasets.
- An algorithm dominates in ROC space only if it dominates in *P*-*R* space (Davis and Goadrich 2006).
- Most significantly, not taking into account N_{TN}. A subjective criterion is necessary to define the length of time that counts as one *TN* value over a continuous recording that contains no targeted vocalizations, whereas N_{TN} does not need to be calculated for the *P*-*R* metric, so *P*-*R* values are better suited to analyzing these time-continuous data.

A.6.4. Signal-to-Noise Ratio

The signal-to-noise ratio (SNR) is the ratio of signal power (P_s) to noise power corrupting the signal (P_n). The SNR compares the level of the desired signal to the level of the background noise; the greater this ratio, the less obtrusive the background noise. SNR is defined in decibels as:

$$\mathrm{SNR} = 10\log_{10}\left(\frac{P_s}{P_n}\right) \tag{2}$$

The signal power of a call in a spectrogram is the average power of the call within the frequency range of the vocalization; the noise power is the average power before and after the call within the same frequency band (Mellinger 2004, Mellinger and Clark 2006). The duration of the noise signal measured before and after the call equals the duration of the call (Figure A-10). This calculation was performed on the original spectrogram without noise reduction. To quantify detector performance for various SNRs, N_{FN} and N_{TP} were calculated for SNR intervals of < 0 dB, 0–5 dB, 5–10 dB, and ≥ 10 dB. *P* values are influenced by the background noise and not by the SNR of the calls. Therefore, *P* values per SNR intervals were not calculated.

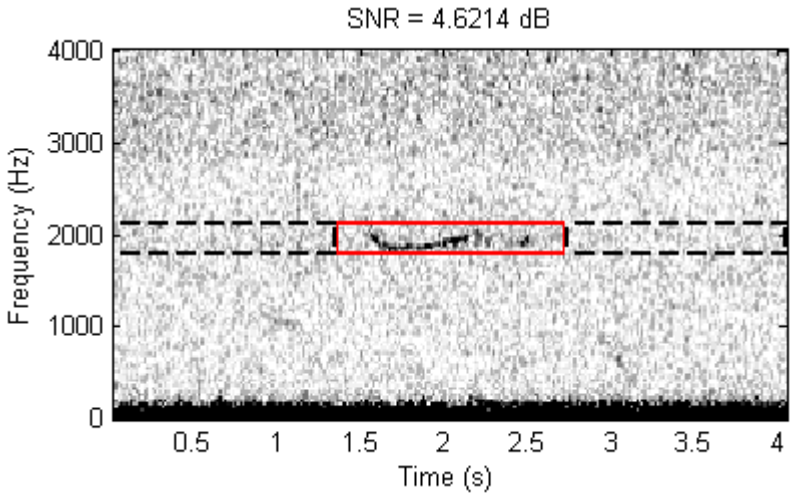

Figure A-10. Calculation of the signal-to-noise ratio (SNR). The power of the call (P_s) is calculated in the red box; the power of the noise (P_n) is calculated in the black boxes on either side of the call.

A.7. Call Count Estimation

Because the detectors/classifiers have false alarms and missed calls, they are imperfect and, as such, the number of automated detections does not exactly equal the actual number of calls present in the recordings. A better estimate can be achieved using *P* and *R*. These values characterize the relationship between the detector/classifier and the dataset. Therefore, these values are specific to, and depend on, both the detector/classifier and the dataset. If the subset of data used to characterize *P* and *R* is representative of the entire dataset, *P* and *R* can be used to extrapolate the total number of vocalizations from the number of detected vocalizations. The total number of detections (N_{det}) found by the detector/classifier is the sum of the number of true and false positives:

$$N_{\text{det}} = N_{TP} + N_{FP} \tag{3}$$

From the definition of *P* (Equation 1), N_{TP} can be defined as:

$$N_{TP} = P \cdot (N_{TP} + N_{FP}) = P \cdot N_{\text{det}} \tag{4}$$

The total number of vocalizations in the data (N_{voc}) is the sum of those correctly identified (*TP*) and those that were missed (*FN*):

$$N_{voc} = N_{TP} + N_{FN} \tag{5}$$

Therefore, *R* becomes:

$$R = \frac{N_{TP}}{N_{TP} + N_{FN}} = \frac{N_{TP}}{N_{voc}} \tag{6}$$

Combining Equations 4 and 6 yields the total number of vocalizations in terms of P, R, and the number of detections:

$$N_{voc} = \frac{N_{TP}}{R} = \frac{P \cdot N_{\text{det}}}{R} \tag{7}$$

All call-count estimation plots in the main report (bubble-plots) were produced using Equation 7.

A.8. Detector/Classifier Performance

The performance of each automated detector/classifier is evaluated using test datasets from both the winter 2013–2014 and summer 2014 programs. The test datasets consist of all fully manually-annotated data samples for each program. For each detector/classifier and each season's dataset, the precision (*P*) and recall (*R*) of the detector/classifier on the entire test dataset are given. The SNR distribution of the test dataset over four SNR intervals and the *R* values calculated for each SNR interval are shown in Figure A-11.

A.8.1. Bowhead Winter Songs

The bowhead winter song detector/classifier was tested against the fully manually-annotated recordings from winter 2013–2014. The test dataset had 475 manually-annotated bowhead songs (Figure A-11, left). The performance of the bowhead song detector/classifier on the test dataset yielded $P = 0.74$ and $R = 0.30$. As expected, the detector/classifier was able to detect more calls at higher SNRs (Figure A-11, right). The highest *R* value was 0.45, obtained for calls with SNRs between 5 and 10 dB.

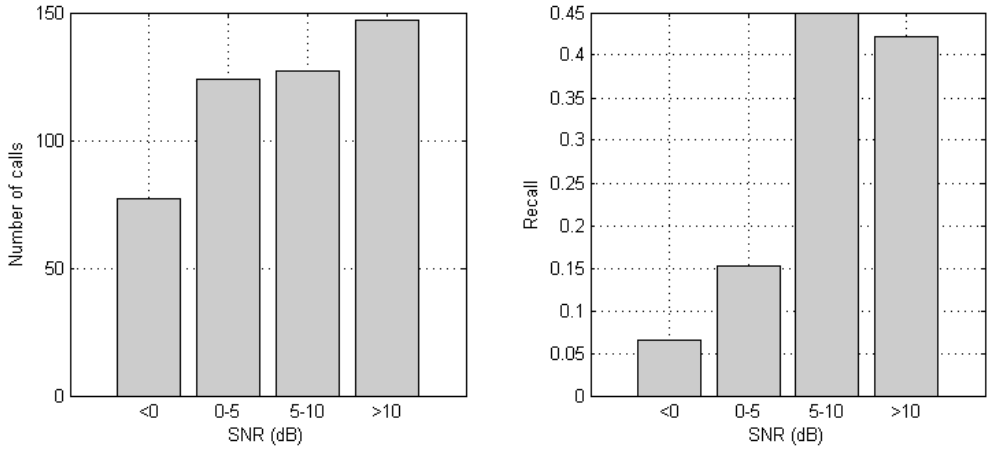

Figure A-11. Performance of the bowhead winter song detector/classifier on the winter 2013–2014 test dataset. (Left) Signal-to-noise ratio (SNR) distribution of calls in the test dataset. (Right) Recall of the detector/classifier per call SNR interval.

A.8.2. Bowhead Summer Moans

The bowhead summer moan detector/classifier was tested against fully-annotated recordings collected during the summer 2014 monitoring program. The test dataset had 527 manually-annotated bowhead moans (Figure A-12, left). The performance of the bowhead moan detector/classifier on the test dataset yielded $P = 0.81$ and $R = 0.39$. As expected, R increased with increasing SNR (Figure A-12, right).

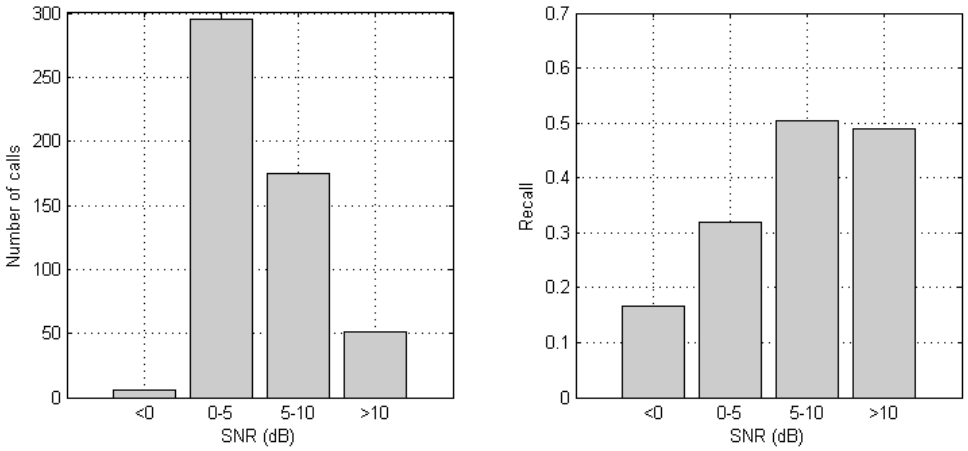

Figure A-12. Performance of the bowhead summer moan detector/classifier on the summer 2014 test dataset. (Left) Signal-to-noise ratio (SNR) distribution of calls in the test dataset. (Right) Recall of the detector/classifier per call SNR interval.

A.8.3. Beluga Whistles

The beluga whistle detector/classifier was used to analyze data from both the winter 2013–2014 and the summer 2014 monitoring programs. Performance of the detector/classifier was evaluated separately on each program given the very different characteristics of the background noise between the summer and the winter.

Winter 2013–2014 Program

The test winter dataset had 497 manually-annotated beluga whistles (Figure A-13, left). Most annotated whistles had a SNR between 0 and 5 dB. The beluga whistle detector/classifier had $P = 0.56$ and $R = 0.55$. The highest R was 0.67, obtained for whistles with SNR > 10 dB (Figure A-13, right).

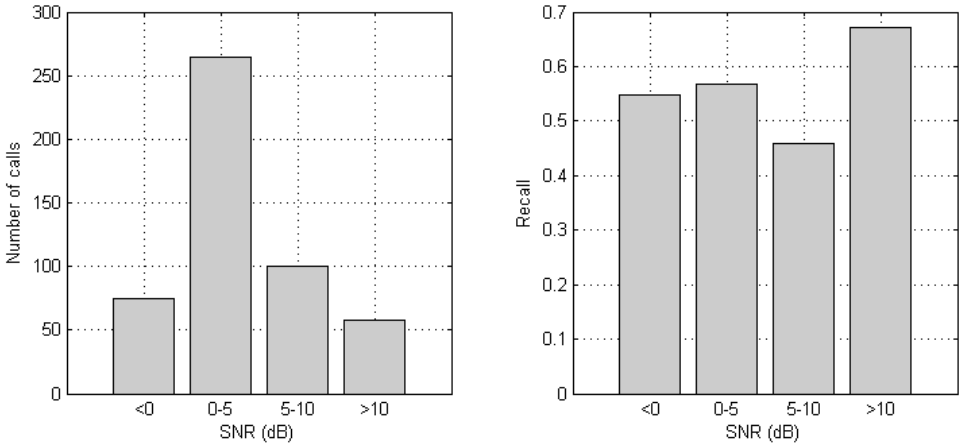

Figure A-13. Performance of the beluga whistle detector/classifier on the winter 2013–2014 test dataset. (Left) Signal-to-noise ratio (SNR) distribution of calls in the test dataset. (Right) Recall of the detector/classifier per call SNR interval.

Summer 2014 Program

The summer test dataset had 496 manually-annotated beluga whistles (Figure A-14 , left). Most annotated whistles had a SNR between 0 and 5 dB. The beluga whistle detector/classifier had $P = 0.68$ and $R = 0.55$. The highest R was 0.57, obtained for whistles with SNR between 0 and 5 dB (Figure A-14, right)

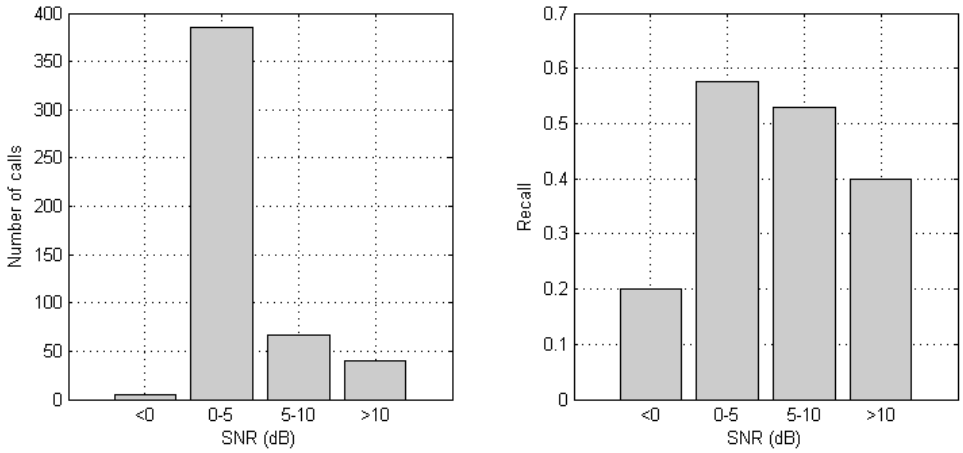

Figure A-14. Performance of the beluga whistle detector/classifier on the summer 2014 test dataset. (Left) Signal-to-noise ratio (SNR) distribution of calls in the test dataset. (Right) Recall of the detector/classifier per call SNR interval.

A.8.4. Walrus Grunts

Walrus grunts were recorded in both winter and summer. Therefore, the performance of the walrus grunt detector/classifier was calculated independently for the summer 2014 and for the winter 2013–2014 datasets (i.e., one set of P and R values for each datasets).

Winter 2013–2014 Program

The winter 2013–2014 test dataset had 231 manually annotated walrus calls (Figure A-15, left). The walrus call detector/classifier had $P = 0.72$ and $R = 0.28$ for the winter 2013–2014 test dataset. R for calls with a SNR > 10 dB is lower than that for calls with a SNR of 5–10 dB due to the misrepresentation of that SNR interval in the winter 2013–2014 test dataset (only 13 walrus grunts annotated with a SNR > 10 dB, (Figure A-15, right). Similarly, R for calls with a SNR < 0 dB is higher than that for calls with a SNR of 0–5 dB due to the under-representation of that SNR interval (only 15 annotated walrus grunts.

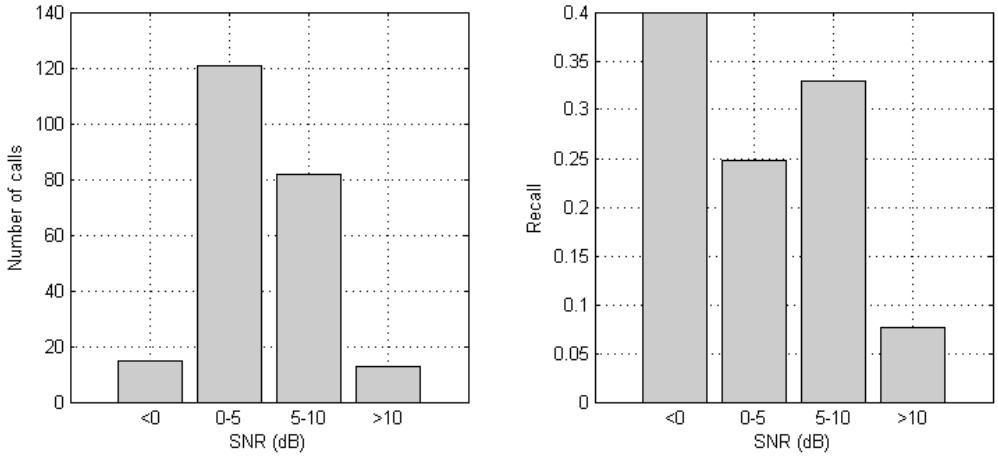

Figure A-15. Performance of the walrus grunt detector/classifier on the winter 2013–2014 test datasets. (Left) Signal-to-noise ratio (SNR) distribution of calls in the combined test datasets. (Right) Recall of the detector per call SNR interval.

Summer 2014 Program

The walrus grunt detector/classifier was tested against fully-annotated recordings collected during the summer 2014 program. The test dataset had a total of 358 manually-annotated walrus grunt (Figure A-16, left). The performance of the walrus grunt detector/classifier on the test dataset yielded $P = 0.52$ and $R = 0.88$. The highest R was 0.53 and was obtained for grunts with SNR between 5 and 10 dB (Figure A-16, right).

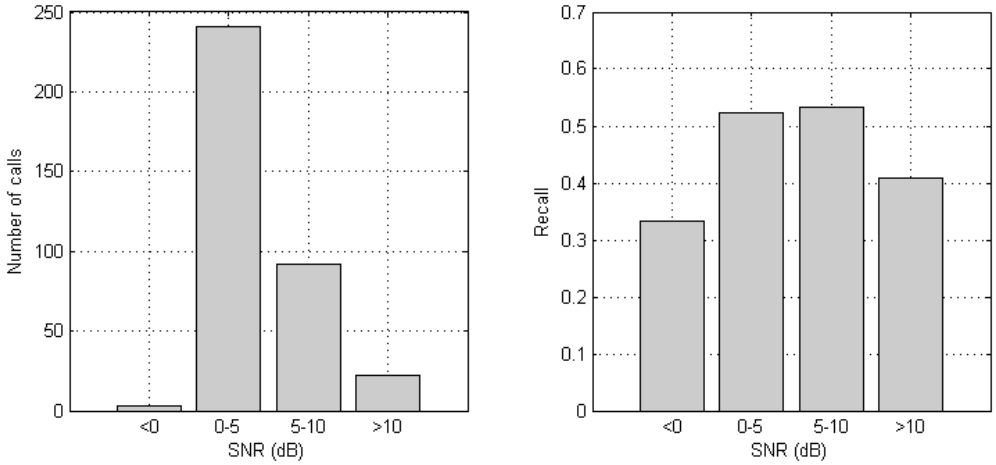

Figure A-16. Performance of the walrus grunt detector/classifier on the summer 2014 test datasets. (Left) Signal-to-noise ratio (SNR) distribution of calls in the combined test datasets. (Right) Recall of the detector per call SNR interval.

A.8.5. Bearded Seal Calls

Bearded seal calls were recorded in both winter and summer. Therefore, the performance of the bearded seal calls detector/classifier was calculated independently for the summer 2014 and for the winter 2013–2014 datasets (i.e., one set of *P* and *R* values for each dataset).

Winter 2013–2014 Program

The winter 2013–2014 test dataset had 653 manually-annotated bearded seal calls (Figure A-17, left). The bearded seal call detector/classifier had $P = 0.77$ and $R = 0.65$ for this dataset. The highest *R* was 0.96 and was obtained for grunts with SNR > 10 dB (Figure A-17, right).

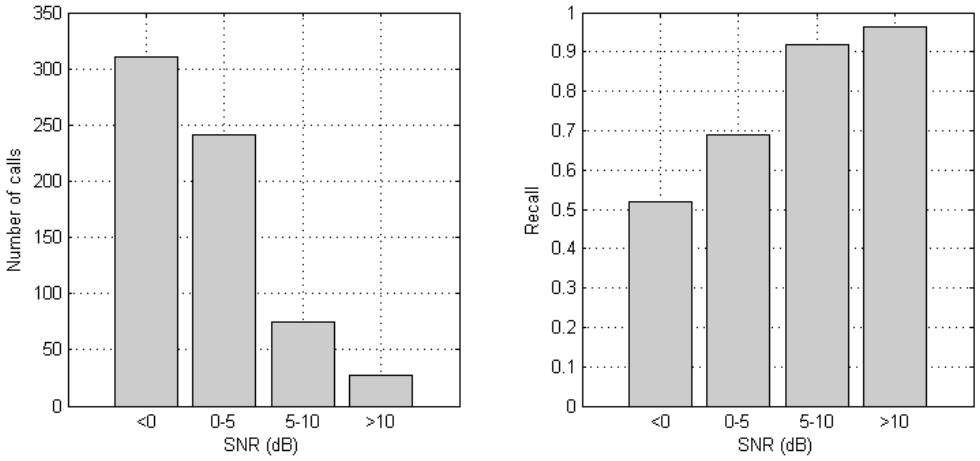

Figure A-17. Performance of the bearded seal detector/classifier on the winter 2013–2014 test dataset. (Left) Signal-to-noise ratio (SNR) distribution of calls in the test dataset. (Right) Recall of the detector/classifier per call SNR interval.

Summer 2014 Program

The summer 2014 test dataset had 379 manually-annotated bearded seal calls (Figure A-17, left). The bearded seal call detector/classifier had $P = 0.84$ and $R = 0.58$ for this dataset. The highest R was 0.78 and was obtained for grunts with SNR > 10 dB (Figure A-17, right).

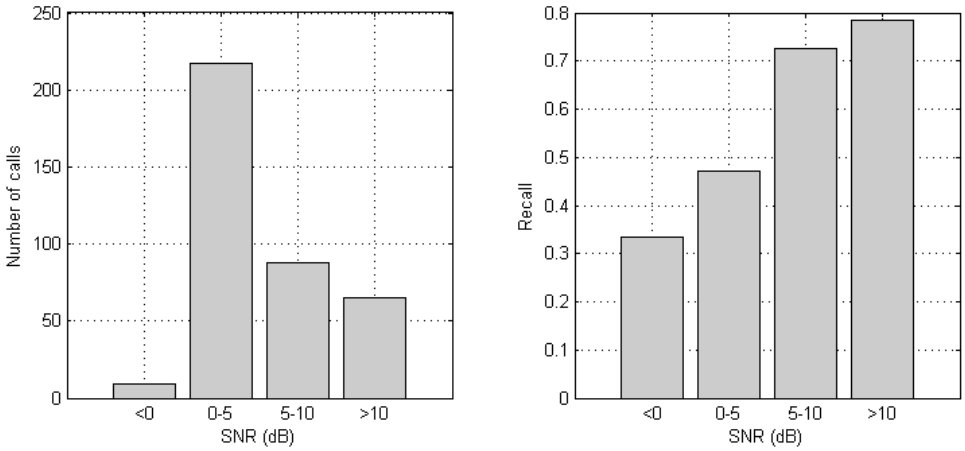

Figure A-18. Performance of the bearded seal detector/classifier on the summer 2014 test dataset. (Left) Signal-to-noise ratio (SNR) distribution of calls in the test dataset. (Right) Recall of the detector/classifier per call SNR interval.

A.8.6. Minke whale boing calls

The minke whale boing call detector was only run on the winter 2013–2014 data. The overwinter 2013 test dataset had 462 manually-annotated minke whale boing call (Figure A-19, left). The minke whale boing call detector had $P = 0.95$ and $R = 0.50$ for this dataset. The highest R was 0.67 and was obtained for boing calls with SNR between 5 and 10 dB (Figure A-19, right).

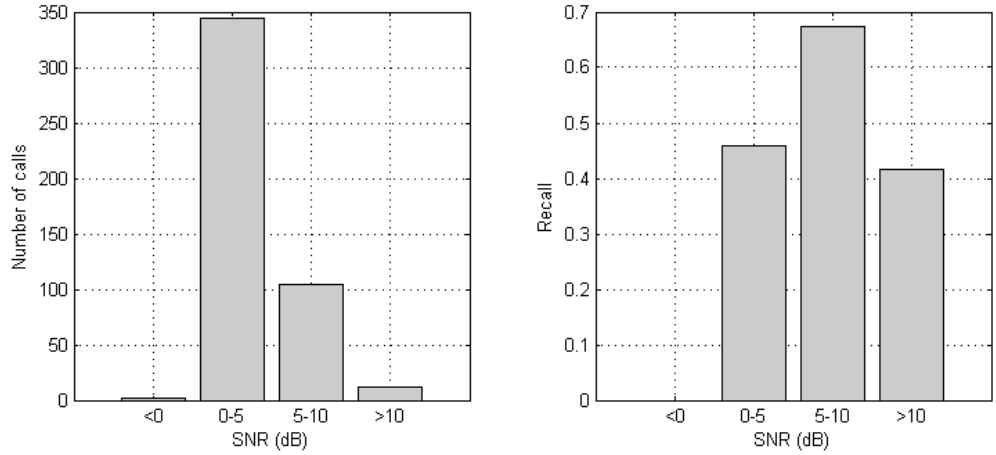

Figure A-19. Performance of the minke whale boing call detector on the winter 2013–2014 test dataset. (Left) Signal-to-noise ratio (SNR) distribution of calls in the test dataset. (Right) Recall of the detector/classifier per call SNR interval.

A.8.7. Summary

Table A-11 presents the summary of the performance of each detector/classifier used in this study for each season.

Table A-11. Precision (*P*) and recall (*R*) for all SNRs of each detector/classifier.

Detector	*P*	*R*
Bowhead songs, winter	0.74	0.30
Bowhead moans, summer	0.81	0.39
Beluga whistles, winter	0.56	0.55
Beluga whistles, summer	0.68	0.55
Walrus grunts, winter	0.72	0.28
Walrus grunts, summer	0.52	0.88
Bearded seal, winter	0.77	0.65
Bearded seal, summer	0.84	0.58
Minke whale, winter	0.95	0.50

A.9. Probability of Detection by Manual Analysis

To determine whether manually reviewing only 5% of the data provided an accurate estimate of the acoustic occurrence of marine mammal calls, analysts randomly selected, then fully-annotated more than 43 h of acoustic data containing a representative sample of the commonly detected species, specifically bowhead, beluga and gray whales, bearded and ringed seals, and walrus. Selected files were distributed across stations and over the whole recording period. For each file, an algorithm written for this purpose then chose a random start time and manually searched the next *n*% of the file (corresponding to the analysis sample) for manual annotations. *n* was varied from 1 to 100% in increments of one. This random sample selection was iterated 2000 times per file for each sample size. A detection probability (DP) was obtained for each file and sample size by dividing the number of samples containing at least one annotation in the random sample by 2000. The comparison of detection probabilities across the sampling period provided an overview of seasonal and inter-specific variations.

A.9.1. Manual Analysis Detection Probability: Winter 2011–2012 and Summer 2012 Programs

Samples of data of 5% of each acoustic data file were manually analyzed to determine the presence of calls from each species in the winter and summer datasets. Calls were separated by call types (Table A-12). The goal of this analysis was to assess and validate the protocol of manual examination of a fraction of the datasets. The 5% manual analysis protocol is compared to estimated 1%, 2%, and 10% manual analysis protocols (Table A-13, Table A-14).

For bowhead whales, analysts annotated individual sounds, but did not distinguish or characterize songs (see for example Delarue et al. 2009). If analysts were unable to definitively identify a species in a sample, they would examine the source file of the sample for more easily identifiable calls within the same time window.

Table A-12. Call types by species annotated during manual analysis of the winter 2012–2013 and summer 2013 datasets. Abbreviations: AM = amplitude-modulated, FM = frequency-modulated, HF = high-frequency, LF = low-frequency.

Species	Call type	Description
Bowhead whale	Upsweep	Upsweeping FM tonal, usually below 600 Hz.
	Downsweep	Downsweeping FM tonal, usually below 600 Hz.
	Constant	Relatively flat FM tonal, usually below 600 Hz.
	Convex	Inflected FM tonal, increasing then decreasing in frequency. Usually below 600 Hz.
	Concave	Inflected FM tonal, decreasing then increasing in frequency. Usually below 600 Hz.
	Complex	FM moans with more than one inflection point and/or with harmonics. Any FM and AM calls extending above 600 Hz.
	Overlap	Overlapping calls produced concurrently by several individuals.
	Other	Bowhead calls outside the above categories.
Walrus	Knock	Broadband impulsive sounds typically occurring in long series.
	Bell	Tonal calls centered around 450 Hz and typically associated with knocks.
	Chimp	Two-part call reminiscent of chimpanzee vocalizations and often produced in long sequences. Sometimes repeated without interruption between consecutive units. Second part higher in frequency than first part.
	Grunt	Grunting sound. Often produced in pairs or triads repeated in long sequences.
	Bark	Often produced in pairs or triads repeated in long sequences. Similar to grunts, but higher in frequency (400 Hz).
	Snort	Snorting/burping sound typically increasing in frequency. Typically not produced in sequence.
	Tone	LF tonal calls, typically flat or downsweeping. Usually around 100–200 Hz. Similar to bowhead moans but shorter (< 0.5 s).
	Low-frequency downsweep	A short call (< 0.5 s) with features intermediate between a grunt and tone; fast downward sweep rate; less than 100 Hz and emphasis on LF (< 50 Hz)
	Overlap	Overlapping calls produced by several animals concurrently.
	Other	Walrus calls outside the above categories.
Beluga whale	Low whistle	FM calls without harmonics below 2500 Hz.
	High whistle	FM calls without harmonics above 2500 Hz.
	Buzz	Broadband buzzing sounds.
	Chirp	Very short, HF sound. Reminiscent of small bird chirps.
	Click	Broadband clicks, presumably echolocation related.
	Overlap	Overlapping calls produced by several animals concurrently.
	Other	Beluga calls outside the above categories.
Bearded seal	Long trill	Downsweeping trills longer than 6 s.
	Short trill	Downsweeping trills shorter than 6 s.
	Upsweeping trill	All upsweeping trills.
	Constant trill	Flat trills.
	Complex trill	Trills containing both up- and downsweeping segments.
	Overlap	Overlapping calls produced by several animals concurrently.

Species	Call type	Description
	Other	Bearded seal calls outside the above categories.
Fin whale	20 Hz pulse	Pulse downsweeping from 25 to 18 Hz, about 1 s long.
	Broadband downsweep	Same bottom frequency as 20 Hz pulse, but top frequency can extend up to 50 Hz or above.
	Other	Calls that do not match the above categories.
Gray whale	Knock	Knocking sounds. No frequency modulation.
	Click	Series of impulsive sounds similar to knocks but varying in pitch throughout the series.
	Grunt-like knock	Superposition of knocks and grunts.
	Moan/growl	Moans with harmonic. Very LF (fundamental near 100 Hz) with growly texture. Sometimes mixed with grunt-like knocks.
	Other	Calls outside the above categories.
Humpback whale	Grunt/snort, wop	AM calls often ascending in frequency at the end (e.g., Thompson et al. 1986).
	Other	Calls outside the above categories (e.g., moans, cries, etc.).
Killer whale	Pulsed call	Characterized by harmonic structure. Fundamental frequency usually around 800–1000 Hz. Expect repetitions of stereotyped calls within files.
	Whistle	FM calls usually without harmonics.
	Other	Calls outside the above categories.
Minke whale	Boing	Pulsed call with fundamental frequencies and harmonics around 1200–1500 Hz, 1–2 s long.
Ribbon seal	Medium downsweep	FM calls, sometimes with harmonic, downsweeping from 2–5 kHz to 100 Hz, usually < 2 s. Metallic texture and sonority.
	Other	Primarily contains the puffing sounds described by Watkins and Ray (1977). Includes other uncategorized calls.
Ringed seal	Bark	Short barking/grunting sounds below 1 kHz and produced in series; often alternating with yelps.
	Yelp	Short yelping sounds between 600–1000 Hz; can occur alone or in mixed sequences with barks.
	Other	Ringed seal calls outside the above categories.
Unknown	Undescribed	Any biological sound that cannot be classified as one of the above species; includes isolated calls that cannot be assigned to a species based on context. Most presumed ice seal calls were likely logged here.
	Grunt	Any grunt-like calls not likely produced by walrus.

The estimated DP for selected files that contain bowhead, beluga and gray whale, ringed and bearded seal, and walrus (Table A-13, Table A-14) calls indicate that the performance of the manual analysis protocol[2] varies with species and season.

Bowhead calls had a mean DP of 0.82 during the winter deployment (range: 0.33 to 1). DPs increased in late October, were highest in November, December, April, and May when bowheads produce long, elaborate songs (Delarue et al. 2009) as they migrate through the Chukchi Sea, and then decreased in late spring. The mean DP in summer 2012 was 0.55, which indicated lower calling rates in summer months. The high DP detected in the file recorded at the end of the

[2] i.e., The probability that a randomly selected 2 min/90 s [winter/summer] sample will contain calls of a given species if calls are present within its 40 min/30 min [winter/summer] source file.

summer program, on 10 Oct, corresponded to the annual increase in vocal activity during fall associated with the onset of singing (Table A-14).

Bearded seal calls had a mean DP of 0.62 (range 0.09 to 1) during the winter deployment. DP was close or equal to 1 from November to early July, with one exception at CL50 in February. The mean DP (0.4) during the summer 2012 deployment persisted into fall, although some peaks in calling activity are possible, as indicated on 11 Oct at CL05 (Table A-13, Table A-14).

Beluga whales' DP was variable (mean: 0.56; range: 0.14 to 1) during the winter deployment. The highest DPs were recorded during the spring migration. The three summer 2012 files analyzed each had DP close or equal to 1 (Table A-13, Table A-14).

Ringed seals' DP was relatively constant throughout the year and consistently low, averaging 0.14 (Table A-13). Although not included in this analysis, summer data follow the same pattern (Delarue et al. 2011). This suggests the current analysis protocol underestimates the presence of ringed seal calls in the data (Table A-13, Table A-14).

Walrus calls typically have a high DP due to high calling rates, with a few exceptions. The mean DP was 0.71 and 0.87 in the winter and summer data, respectively (Table A-13, Table A-14).

Gray whale DP averaged 0.42 in the summer data (range: 0.13–1). A strong variability in DP, and therefore calling rate, was observed (Table A-14).

Table A-13. Manual analysis detection probabilities (DPs) of bowheads, belugas, ringed seals, bearded seals, and walrus for files recorded at several stations during the winter 2011–2012 program based on a manual review of 5% of the data. Results for each species are ordered chronologically. 1%, 2%, and 10% manual analysis protocols are estimates and are included here for comparison.

Species	Station	Date	DP (5%)	DP (1%)	DP (2%)	DP (10%)
Bearded seal	WN60	15 Sep 2011	0.21	0.06	0.10	0.34
	PLN40	17 Oct 2011	0.12	0.04	0.05	0.24
	WN20	16 Nov 2011	0.09	0.02	0.05	0.14
	PLN100	15 Dec 2011	0.94	0.61	0.74	1.00
	PL50	15 Jan 2012	1.00	1.00	1.00	1.00
	CL50	17 Feb 2012	0.24	0.12	0.16	0.27
	B05	19 Mar 2012	0.99	0.76	0.92	1.00
	W35	17 Apr 2012	1.00	1.00	1.00	1.00
	PLN100	3 Jul 2012	1.00	1.00	1.00	1.00
Beluga whale	WN60	7 Oct 2011	0.59	0.22	0.37	0.87
	W35	19 Oct 2011	0.30	0.06	0.13	0.40
	PLN80	7 Nov 2011	0.51	0.16	0.27	0.73
	CL50	23 Nov 2011	0.49	0.17	0.28	0.67
	B05	8 Dec 2011	0.14	0.03	0.06	0.24
	B05	15 Apr 2012	0.41	0.11	0.18	0.65
	PLN40	1 May 2012	0.24	0.07	0.12	0.37
	PL50	16 May 2012	1.00	0.94	0.99	1.00
	B05	1 Jun 2012	1.00	0.86	0.98	1.00
	B05	3 Jul 2012	0.49	0.20	0.31	0.65
	B05	30 Jul 2012	0.98	0.58	0.76	1.00
Bowhead whale	WN80	27 Aug 2011	0.47	0.18	0.27	0.58
	PLN100	9 Oct 2011	0.67	0.20	0.35	0.81
	PLN40	25 Oct 2011	0.92	0.58	0.74	1.00
	W50	4 Nov 2011	1.00	1.00	1.00	1.00
	PL50	17 Nov 2011	1.00	1.00	1.00	1.00
	CL50	15 Dec 2011	1.00	1.00	1.00	1.00
	B05	15 Apr 2012	1.00	0.84	0.95	1.00
	W35	30 Apr 2012	0.81	0.40	0.59	0.93
	PL50	17 May 2012	0.90	0.36	0.61	1.00
	B05	23 May 2012	0.98	0.77	0.90	1.00
	PBN40	10 Jun 2012	0.33	0.10	0.18	0.50
	B05	22 Jul 2012	0.75	0.24	0.40	0.96
Ringed seal	WN60	8 Oct 2011	0.25	0.07	0.12	0.51
	W35	22 Nov 2011	0.40	0.12	0.18	0.55
	PBN20	15 Dec 2011	0.05	0.01	0.03	0.06
	W50	15 Jan 2012	0.15	0.03	0.07	0.31
	PBN40	18 Feb 2012	0.01	0.01	0.01	0.01
	WN80	18 Mar 2012	0.09	0.01	0.03	0.16
	PL50	22 Apr 2012	0.11	0.04	0.08	0.17
	PLN120	13 May 2012	0.08	0.02	0.03	0.09
Walrus	WN80	28 Aug 2011	0.90	0.51	0.68	1.00
	PN120	16 Sep 2011	0.36	0.15	0.24	0.49
	WN20	15 Oct 2011	0.98	0.72	0.84	1.00
	WN40	1 Nov 2011	0.15	0.05	0.09	0.27
	PLN80	3 Dec 2011	0.98	0.73	0.85	1.00
	PLN100	24 Jun 2012	1.00	0.77	0.90	1.00
	B05	26 Jul 2012	0.59	0.29	0.39	0.77

Table A-14. Manual analysis detection probabilities (DPs) of bowheads, belugas, ringed seals, bearded seals, and walrus for files recorded at several stations during the summer 2012 program based on a manual review of 5% of the data. Results for each species are ordered chronologically. 1%, 2%, and 10% manual analysis protocols are estimates and are included here for comparison.

Species	Station	Date	DP (5%)	DP (1%)	DP (2%)	DP (10%)
Bearded seal	CLN90	13 Aug 2012	0.56	0.23	0.34	0.74
	W35	25 Aug 2012	0.02	0.01	0.02	0.02
	B05	10 Sep 2012	0.03	0.03	0.04	0.03
	PLN80	29 Sep 2012	0.40	0.11	0.21	0.61
	CL05	11 Oct 2012	1.00	0.34	0.66	1.00
Beluga whale	B05	15 Aug 2012	0.97	0.38	0.69	1.00
	B50	30 Sep 2012	1.00	1.00	1.00	1.00
	WN40	7 Oct 2012	0.96	0.37	0.63	1.00
Bowhead whale	CLN120	13 Aug 2012	0.45	0.13	0.22	0.67
	W20	9 Sep 2012	0.58	0.15	0.26	0.87
	B30	20 Sep 2012	0.54	0.16	0.30	0.80
	PLN60	30 Sep 2012	0.19	0.05	0.08	0.26
	BG07	10 Oct 2012	0.98	0.44	0.69	1.00
Gray whale	PL50	13 Aug 2012	1.00	1.00	1.00	1.00
	PLN40	26 Aug 2012	0.25	0.05	0.11	0.45
	W50	7 Sep 2012	0.13	0.03	0.05	0.23
	W20	20 Sep 2012	0.57	0.17	0.28	0.77
	PL35	9 Oct 2012	0.15	0.03	0.06	0.28
Walrus	CLN120	10 Aug 2012	1.00	1.00	1.00	1.00
	PLN40	25 Aug 2012	0.36	0.10	0.19	0.61
	PL50	10 Sep 2012	1.00	1.00	1.00	1.00
	WN40	25 Sep 2012	1.00	1.00	1.00	1.00
	BG08	10 Oct 2012	0.98	0.34	0.64	1.00

Figure A-20 suggests that a substantial increase in the length of the analysis sample would be required to reach 50% DP for ringed seals. Bowhead, bearded seal, and walrus DPs with a 5% analysis sample are all above 60% and would not significantly benefit by increasing the sample size. Simply doubling the sample size would raise the DP near or above 70% for all species except ringed seals.

For the summer data, a doubling of the sample size would raise the detection probability of gray whales and bearded seals near 50%, and that of bowheads to just above 70% (Figure A-21).

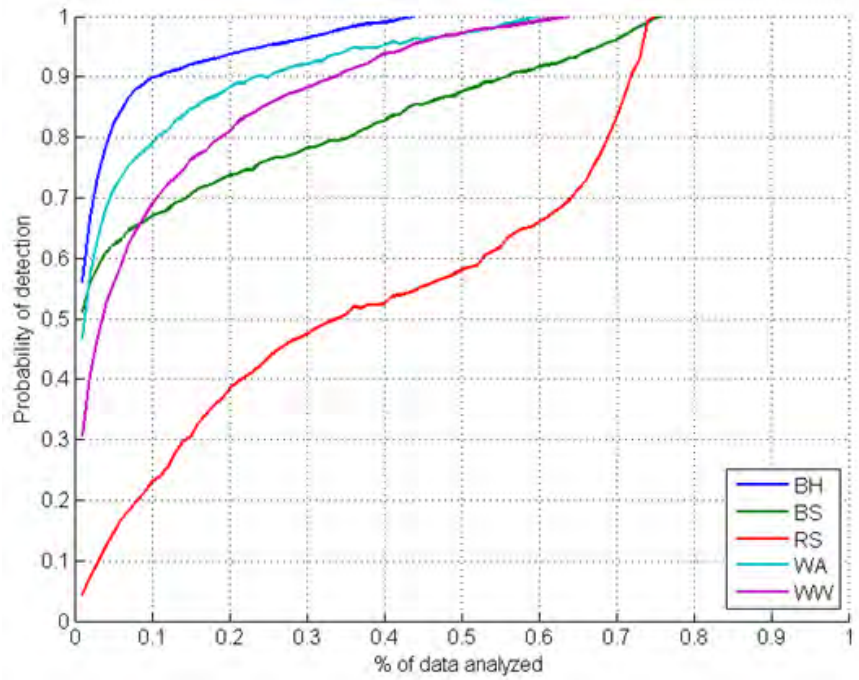

Figure A-20. Detection probability for bowhead and beluga whales, ringed and bearded seals, and walrus as a function of the percent data manually analyzed for a sample of files recorded during the winter 2011-2012 in the northeastern Chukchi Sea.

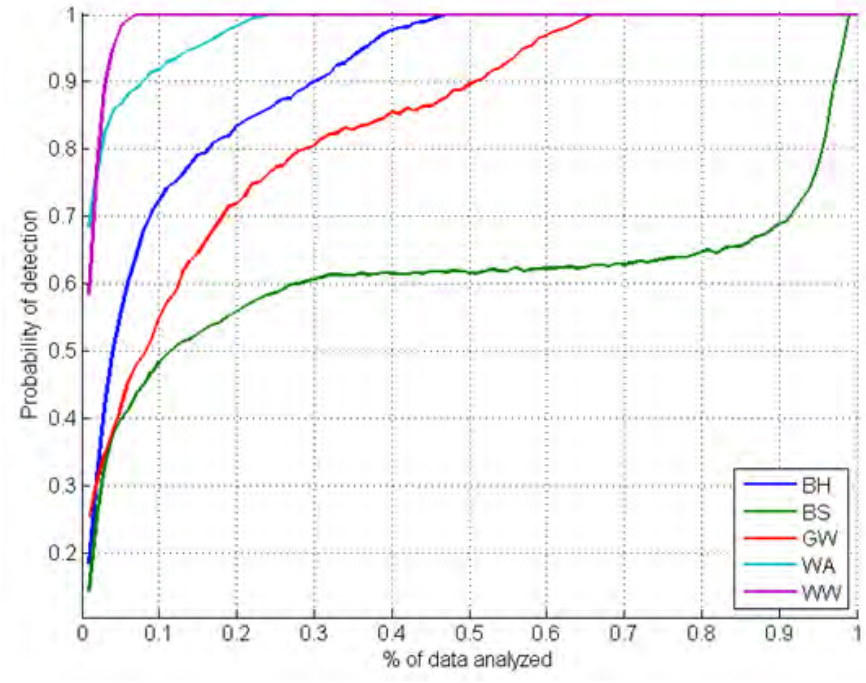

Figure A-21. Detection probability for bowhead, gray, and beluga whales, bearded seals, and walrus as a function of the percent data manually analyzed for a sample of files recorded during the summer 2012 in the northeastern Chukchi Sea.

A.10. Vessel Noise Detection

Ships' propulsion system and other rotating machinery produce narrowband tones (tonals) which can be easily detected. The tonals detector is based on overlapped Fast Fourier Transforms (FFT). The number of seconds of data input to the FFT determines its spectral resolution. Arveson and Vendittis (2000) used both 0.5 and 0.125 Hz resolutions. For this study, spectral analysis was performed at 0.125 Hz resolution by using 8 s of real data with a 2 s advance. This frequency resolution separates the tones from each other for easy detection, and the 2 s advance provides suitable temporal resolution. Higher frequency resolutions can reduce detectability of shipping tones, which are often unstable within 1/16 Hz for long periods. A 120 s long spectrogram is created with 0.125 Hz frequency resolution and 2 s time resolution (524 288-point FFTs, 512 000 real data points, 128 000-point advance, Hamming window). A split-window normalizer (Struzinski and Lowe 1984) distinguishes the tonal peaks from the background (2 Hz window, 0.75 Hz notch, and detection threshold of 4 times the median). The peaks are joined with a 3 × 3 kernel to create contours. Associations in frequency are made if contours occur at the same time. The event time and number of tones for any event at least 20 s long and 40 Hz in bandwidth are recorded for further analysis.

A.11. Seismic Survey Detection

There are measures taken to minimize the occurrence of false alarms, especially from biological sources. For example, sequences with a duration standard deviation greater than 0.2+(number of pulses)/30 s are rejected.

The 100% SEL is computed by adding 0.46 dB to the SEL computed over the 90% rms SPL window, and the pulse time, duration, 90% rms SPL, and SEL are stored for later use. The detected peaks are removed from the event time series and the process is repeated to look for weaker sequences or changes in sequence timing.

This detector requires post-processing to handle some situations. If the pulse period is unstable by more than 0.25 s, the detector cannot "lock-on". Also, if fewer than six pulses occur at the beginning or end of a WAV file at a particular repetition rate, they are missed. Post-processing is applied to address these issues and smooth the results. If at least 8 out of 20 min have seismic detections, then the other 12 min might have missed detecting some seismic pulses. There are three tests to detect possible seismic pulses: (1) The standard deviation in the number of shots per minute is less than 2; (2) The rms SPL during that period is stable within 3 dB and is greater than 125 dB; (3) The 1 min seismic SEL for the minutes with seismic pulses is within 6 dB of the total 1 min SEL. Seismic survey noise is declared missing for the minutes that meet these criteria. The missing minutes are filled in using the 1 min rms SPL and SEL from the ambient computations minus the mean difference between the 1 min seismic SEL and the 1 min ambient SEL.

A.12. Notes on Spectrogram Processing

This report contains many grayscale and color spectrograms representing the spectral evolution with time of sounds recorded during the acoustics programs in the northeastern Chukchi Sea.

The horizontal axis of these figures is time and the vertical axis is frequency, so that the plot provides a visualization of time-varying frequency content of the acoustic data. The spectrograms were processed to exploit the spectral contrast of the signal of interest visually for purposes of the discussion, and therefore the displayed traces do not provide a direct measure of the received SPL.

The caption of each spectrogram describes how the spectrogram was created, including:

FFT Size

Number of points (pts) in each fast Fourier transform (FFT). The acoustic data have a sample rate of 16,384 Hz (samples per second), so a 4096 pt FFT has 4 Hz resolution, and a 16,384 pt FFT has 1 Hz resolution.

Frame Size

Number of actual data points in each FFT. Often less than the FFT size. The actual data points are zero-padded out to the FFT size, which allows display of the spectral content at a high frequency-resolution while maintaining sufficient time resolution for short-duration events. Since many signals of interest are short duration transients, fewer real data points were used in the FFT window to more clearly show the rapid time evolution.

Time step

Number of data points overlapped from one FFT to the next. Generally half the number of real samples, but could be more for finer time resolution.

Window

Type of windowing function applied to the data before FFT to reduce spectral leakage.

Normalization

Most spectrograms in this report are normalized for improved display. Because normalization optimizes contrast in each region of the plot so that both weak and intense signals are similarly visible, the displayed grayscales or colors no longer represent the sound spectral pressure level, as they would without normalization.

Normalization steps we applied:

Step 1: For each frequency bin, compute the average level over the entire file.

Step 2: For each time step, compute a moving average of the results from Step 1, with a frequency bandwidth of 200 Hz.

Normalize each time-frequency bin by the average of Step 1, and the value from Step 2 that is 300 Hz above the current frequency.

Appendix B. Ambient Noise Results

B.1. Analysis Methods

Ambient noise levels at all winter and summer recording stations were examined to document baseline underwater sound conditions in the Chukchi Sea.

Ambient noise levels at each recording station are presented as:

- Statistical distribution of sound pressure levels in each 1/3-octave-band. The boxes of the statistical distributions indicate the first (25%), second (50%), and third (75%) quartiles. The whiskers indicate the maximum and minimum range of the data. The red line indicates the root-mean-square (rms) in each 1/3-octave.
- Spectral level percentiles: Histograms of each frequency bin per 1 min of data. The 5th, 25th, 50th, 75th, and 95th percentiles are plotted. The 95th percentile curve is the frequency-dependent level exceeded by 5% of the 1 min averages. Equivalently, 95% of the 1 min spectral levels are below the 95th percentile curve.
- Broadband and approximate-decade-band sound pressure levels (SPLs) over time for these frequency bands: 10 Hz to 8 kHz, 10–100 Hz, 100 Hz to 1 kHz, and 1–8 kHz.
- Spectrograms: Ambient noise at each station was analyzed by Hamming-windowed fast Fourier transforms (FFTs), with 1 Hz resolution and 50% window overlap. The 120 FFTs performed with these settings are averaged to yield 1 min average spectra.
- Daily cumulative sound exposure levels (SEL (24 h)): computed for the total received sound energy, the detected seismic survey energy, and the detected shipping energy. The SEL (24 h) is the linear sum of the 1 min sound exposure levels (SELs). For shipping, the 1 min SELs (24 h) are the linear 1 min squared rms levels multiplied by the duration, 60 s. For seismic survey pulses, the 1 min SEL is the linear sum of the per-pulse SELs.

The 50th percentile (median of 1 min spectral averages) can be compared to the well-known Wenz ambient noise curves, which show ranges of variability of ambient spectral levels as a function of frequency of measurements off the US Pacific coast over a range of weather, vessel traffic, and geologic conditions. The Wenz curve levels are generalized and are used for approximate comparisons only.

The 1 min averaged, 1 Hz spectral density levels are summed over the 1/3-octave and decade bands to obtain 1 min averaged broadband levels (dB re 1 µPa). These values are available on request. Table B-1 lists the 1/3-octave-band frequencies, Table B-2 the decade-band frequencies. Weather and ice coverage conditions throughout the deployment periods are also provided.

Table B-1. Third-octave-band frequencies (Hz).

Band	Lower frequency	Nominal center frequency	Upper frequency
1	8.9	10	11.2
2	11.6	13	14.6
3	14.3	16	17.9
4	17.8	20	22.4
5	22.3	25	28.0
6	28.5	32	35.9
7	35.6	40	44.9
8	45.0	51	57.2
9	57.0	64	71.8
10	72.0	81	90.9
11	90.9	102	114.4
12	114.1	128	143.7
13	143.4	161	180.7
14	180.8	203	227.9
15	228.0	256	287.4
16	287.7	323	362.6
17	362.7	406	455.7
18	456.1	512	574.7
19	574.6	645	723.9
20	724.2	813	912.6
21	912.3	1024	1149
22	1,150	1,290	1,447
23	1,448	1,625	1,824
24	1,824	2,048	2,297
25	2,298	2,580	2,896
26	2,896	3,251	3,649
27	3,649	4,096	4,597
28	4,598	5,161	5,793
29	5,793	6,502	7,298
30	7,298	8,192	9,195
31	9,195	10,321	11,585
32	11,585	13,004	14,597

Table B-2. Decade-band frequencies (Hz).

Decade-band	Lower frequency	Nominal center frequency	Upper frequency
2	10	50	100
3	100	500	1,000
4	1,000	5,000	10,000

B.2. Winter 2013–2014 Program

B.2.1. One-Third-Octave-Band Sound Pressure and Power Spectral Density Levels

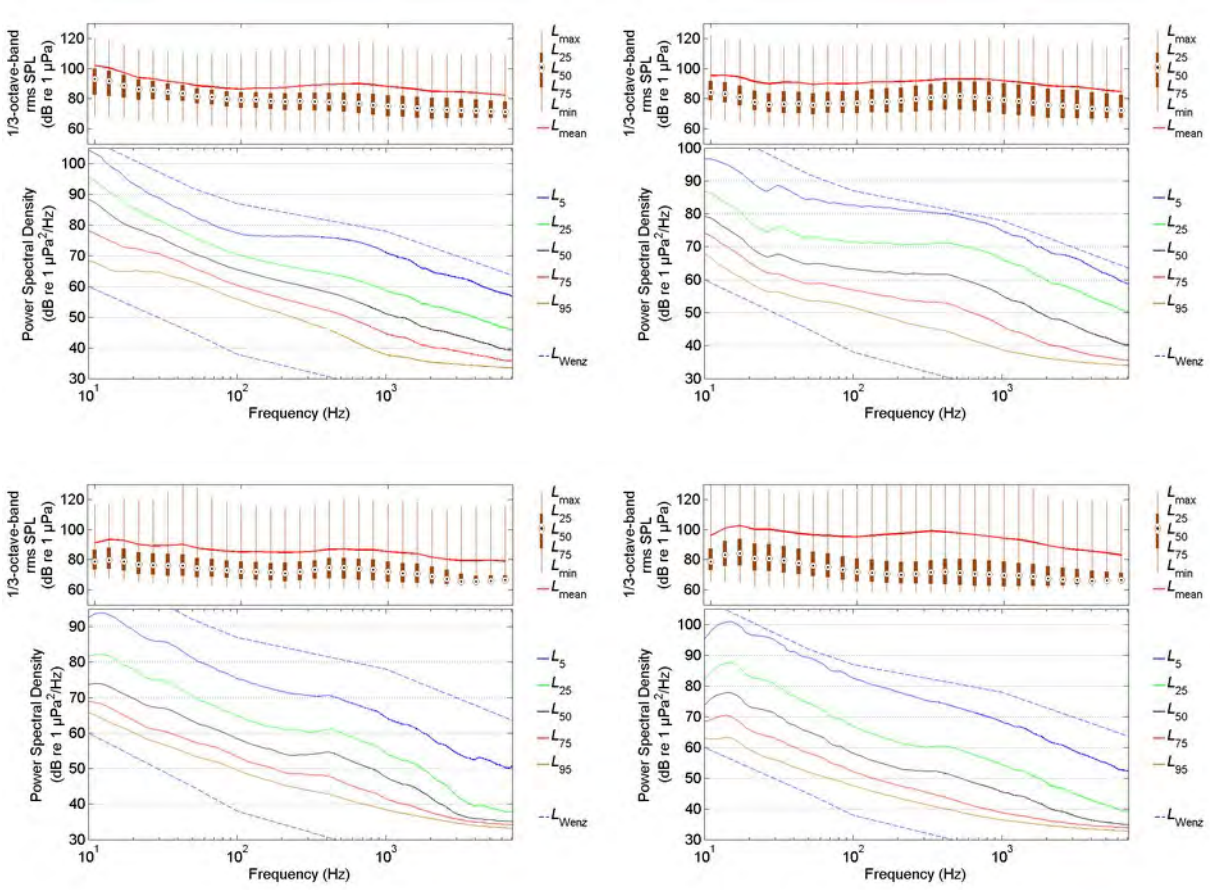

Figure B-1. 1/3-octave-band sound pressure levels and percentile 1 min power spectral density levels for winter 2013–2014 stations. (Top left) B5, (top right) W10, (bottom left) W50, and (bottom right) WN40.

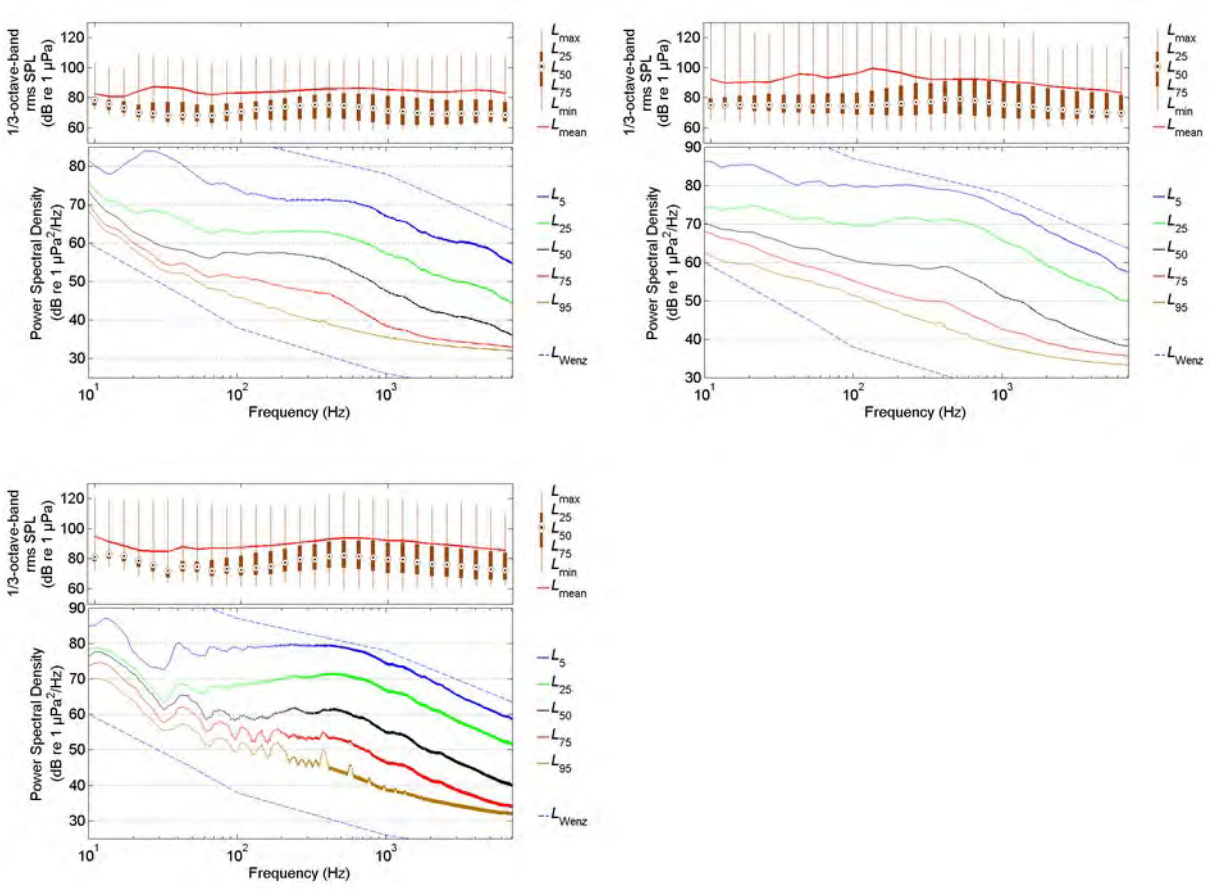

Figure B-2. 1/3-octave-band sound pressure levels and percentile 1 min power spectral density levels for winter 2013–2014 stations. (Top left) PL10, (top right) PL50, and (bottom) CL5.

B.2.2. Broadband and Decade-Band Sound Pressure Levels and Spectrograms

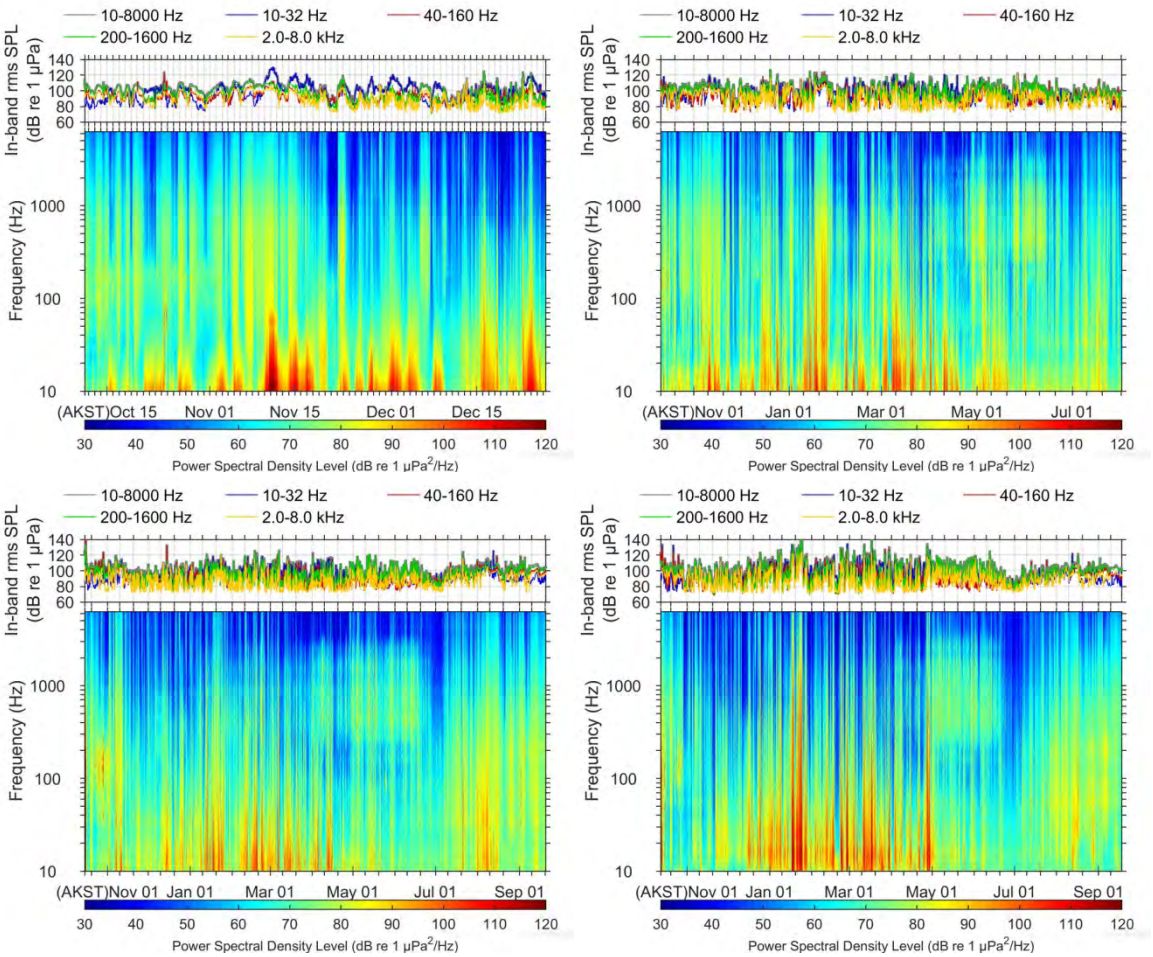

Figure B-3. Broadband and in-band sound pressure levels (SPLs) and spectrograms for winter 2013–2014 stations. (Top left) B5, (top right) W10, (bottom left) W50, and (bottom right) WN40.

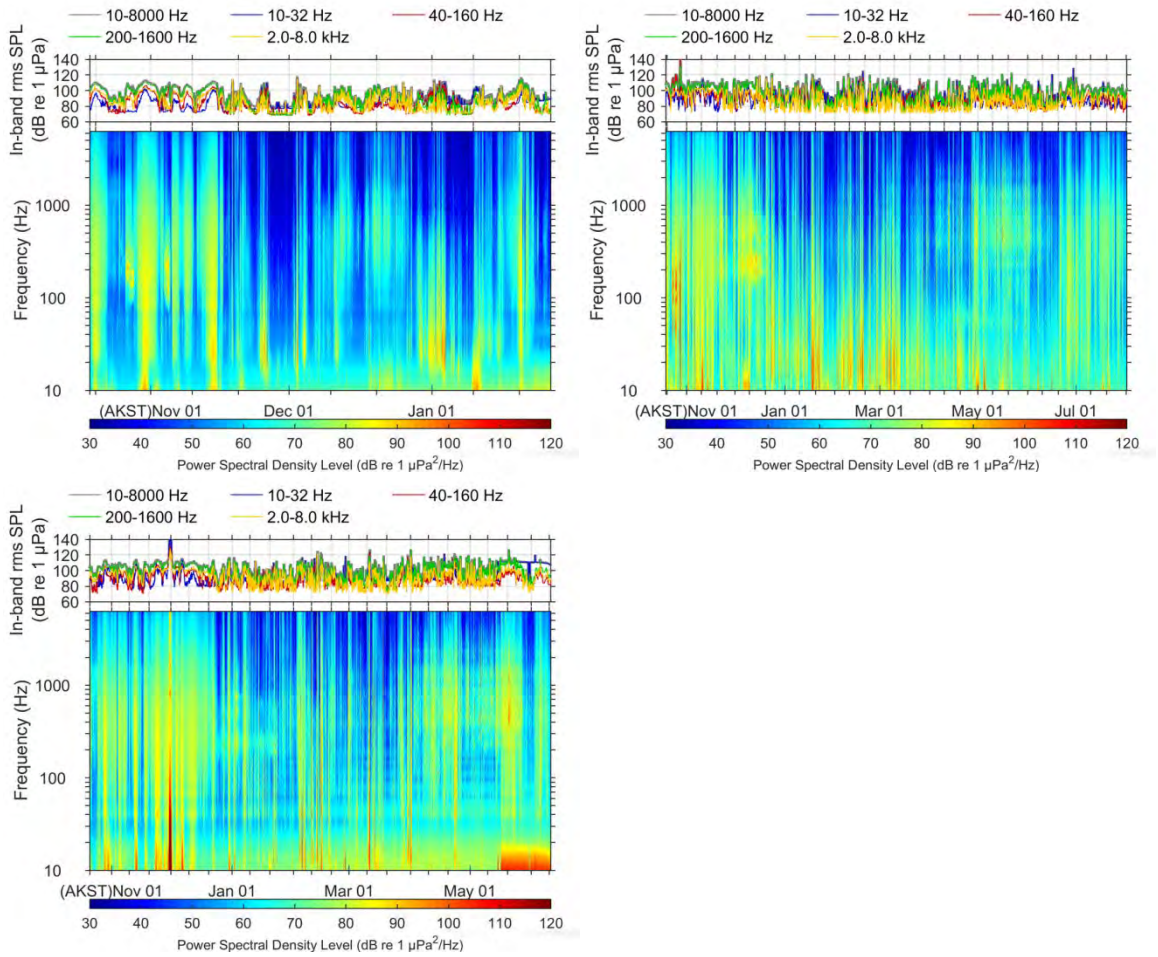

Figure B-4. Broadband and in-band sound pressure levels (SPLs) and spectrograms for winter 2013–2014 stations. (Top left) PL10, (top right) PL50, and (bottom) CL5.

B.2.3. Daily Sound Exposure Level

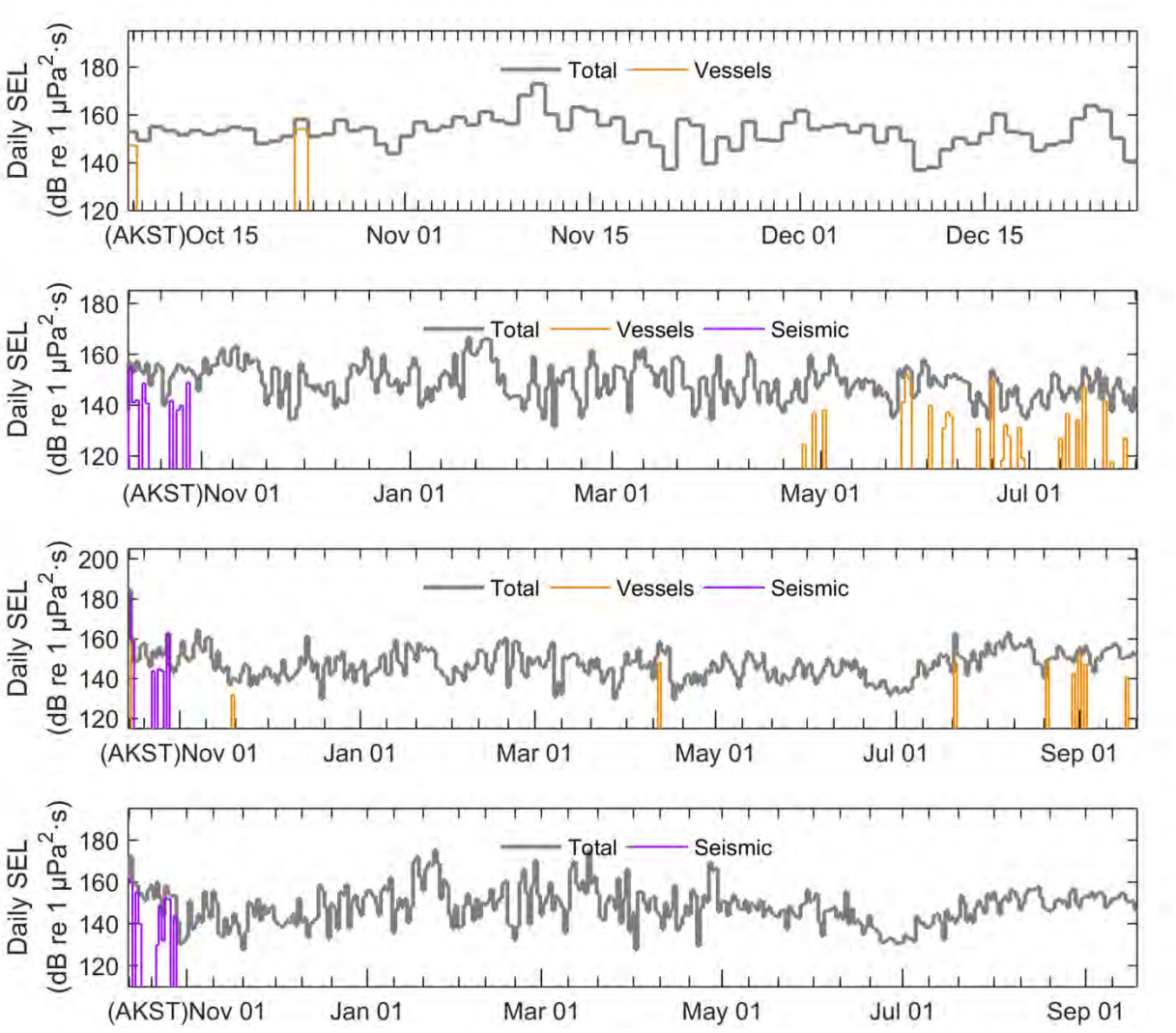

Figure B-5. Daily sound exposure levels (SEL 24 h) for winter 2013–2014 stations. From top to bottom: B5, W10, W50, and WN40.

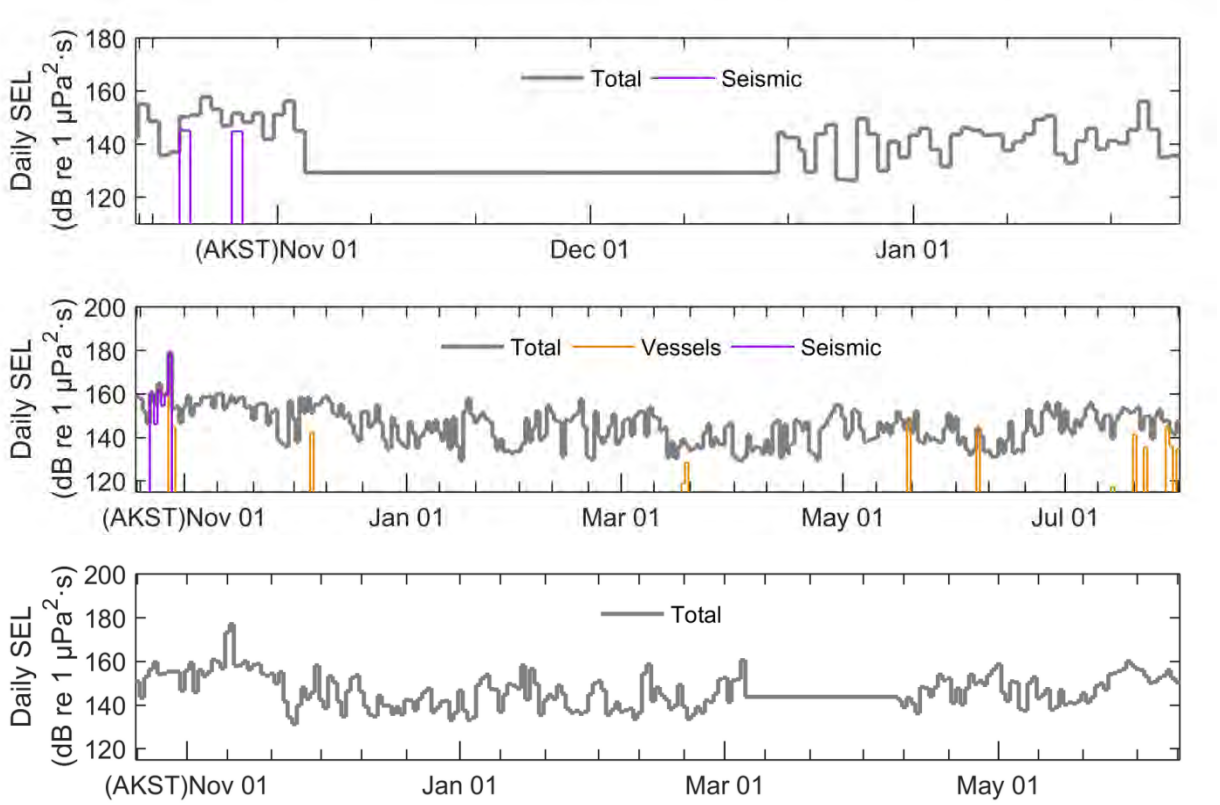

Figure B-6. Daily sound exposure levels (SEL 24 h) for winter 2013–2014 stations. From top to bottom: PL10, PL50, and CL5.

B.3. Summer 2014 Program

B.3.1. One-Third-Octave-Band Sound Pressure and Power Spectral Density Levels

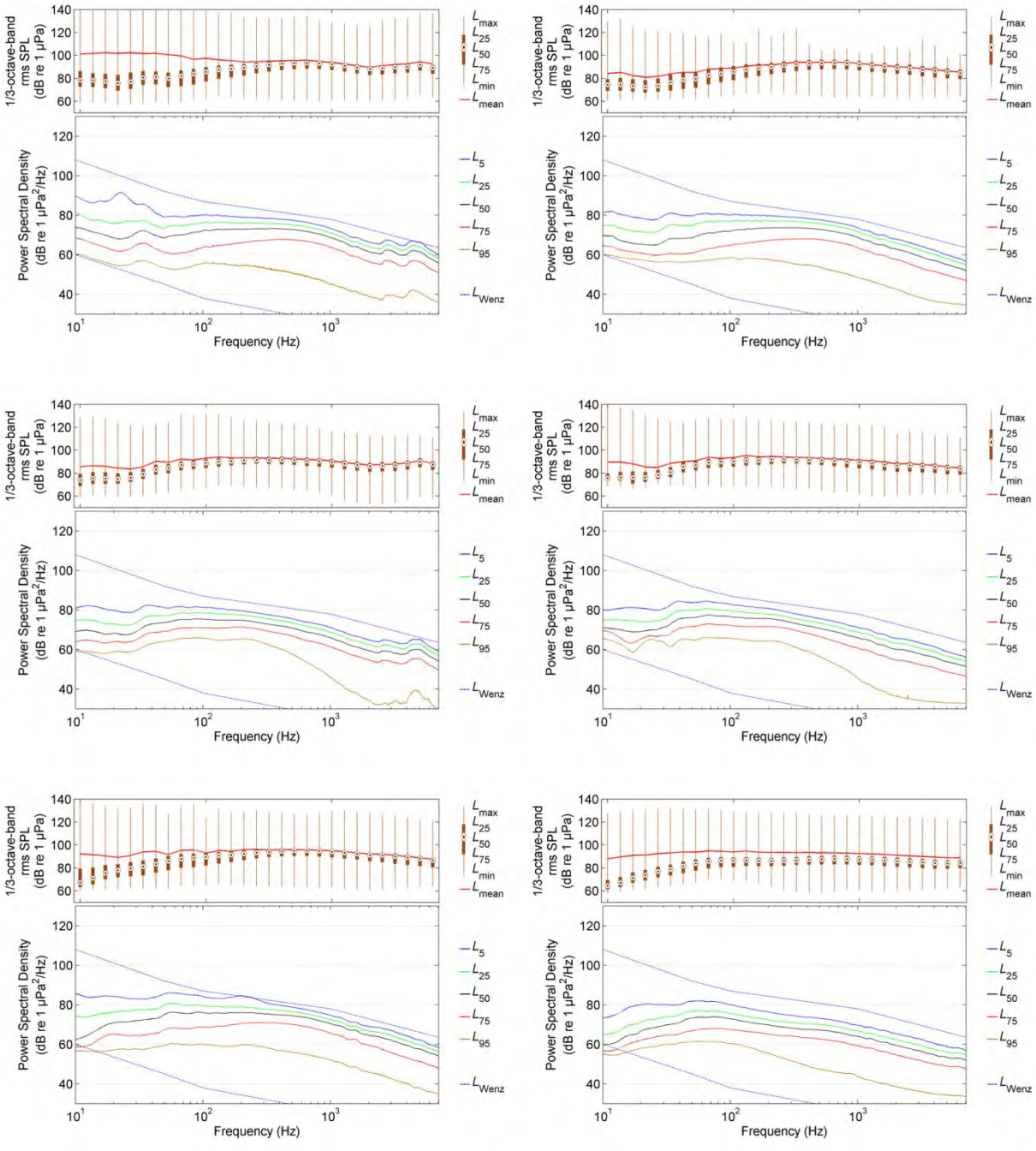

Figure B-7. 1/3-octave-band sound pressure levels and percentile 1 min power spectral density levels for summer 2014 stations. (Top left) W10, (top right) W30, (middle left) W50, (middle right) WN20, (bottom left) WN40, and (bottom right) WN80.

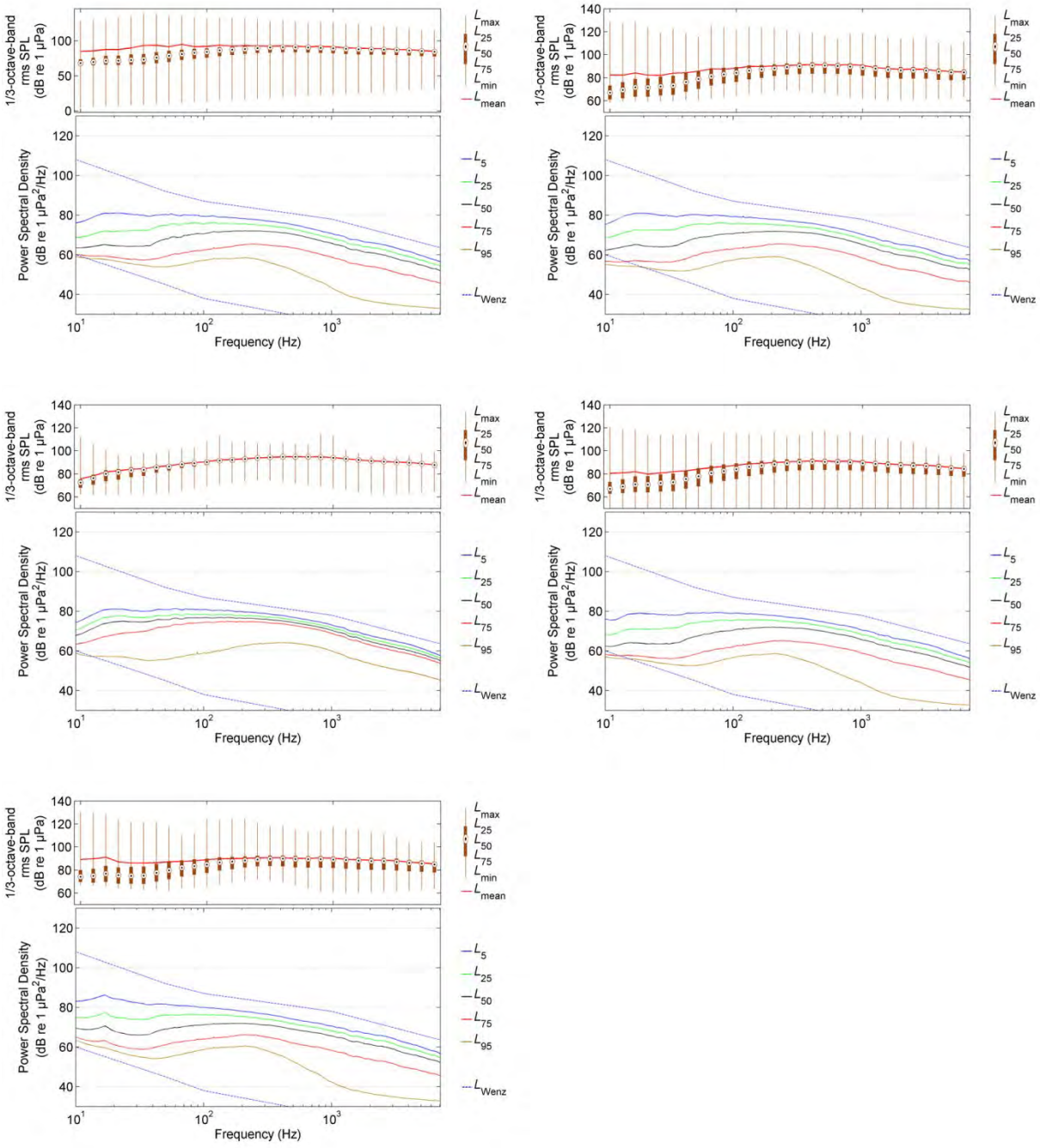

Figure B-8. 1/3-octave-band sound pressure levels and percentile 1 min power spectral density levels for summer 2014 stations. (Top left) BGB, (top right) BGC, (middle left) BGD, (middle right) BGE, and (bottom) BGF.

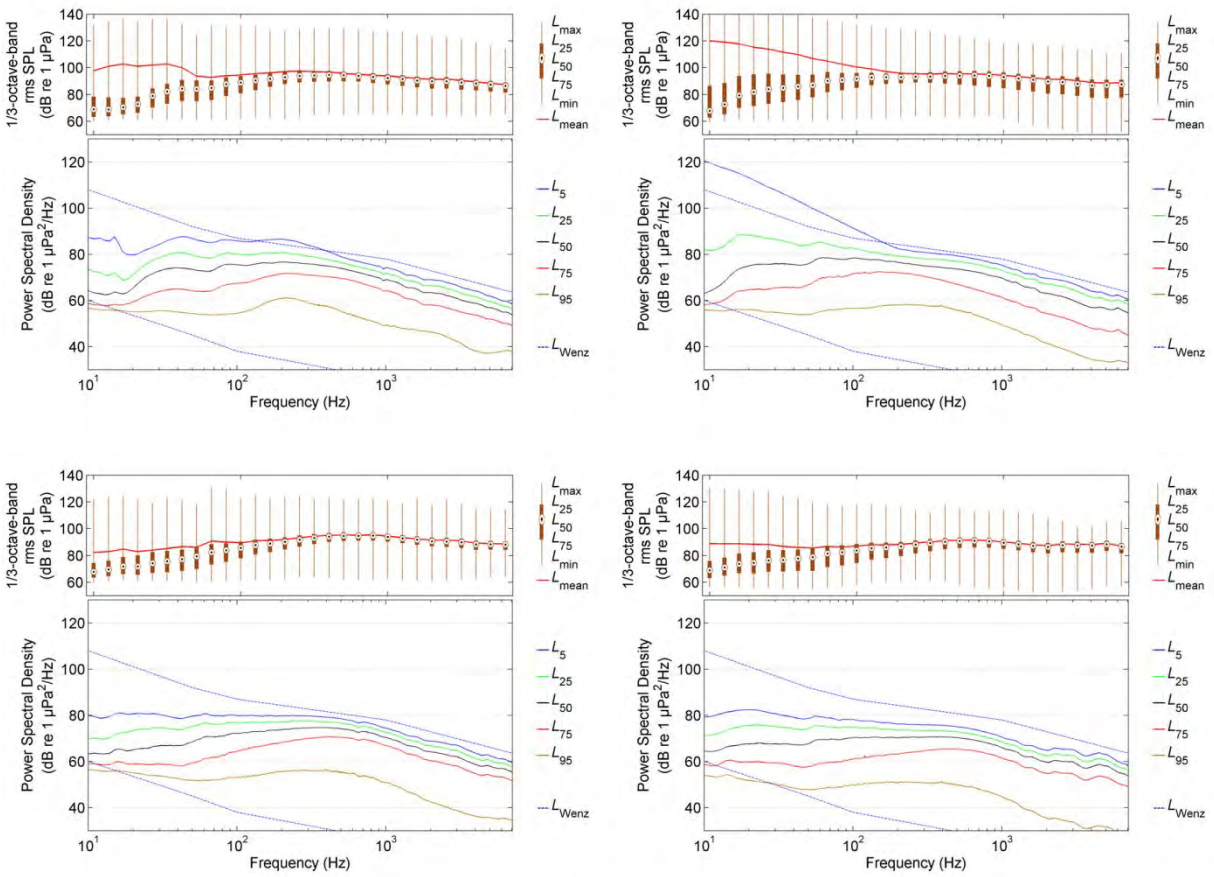

Figure B-9. 1/3-octave-band sound pressure levels and percentile 1 min power spectral density levels for summer 2014 stations. (Top left) PL10, (top right) PL30, (bottom left) PL50, and (bottom right) PLN20.

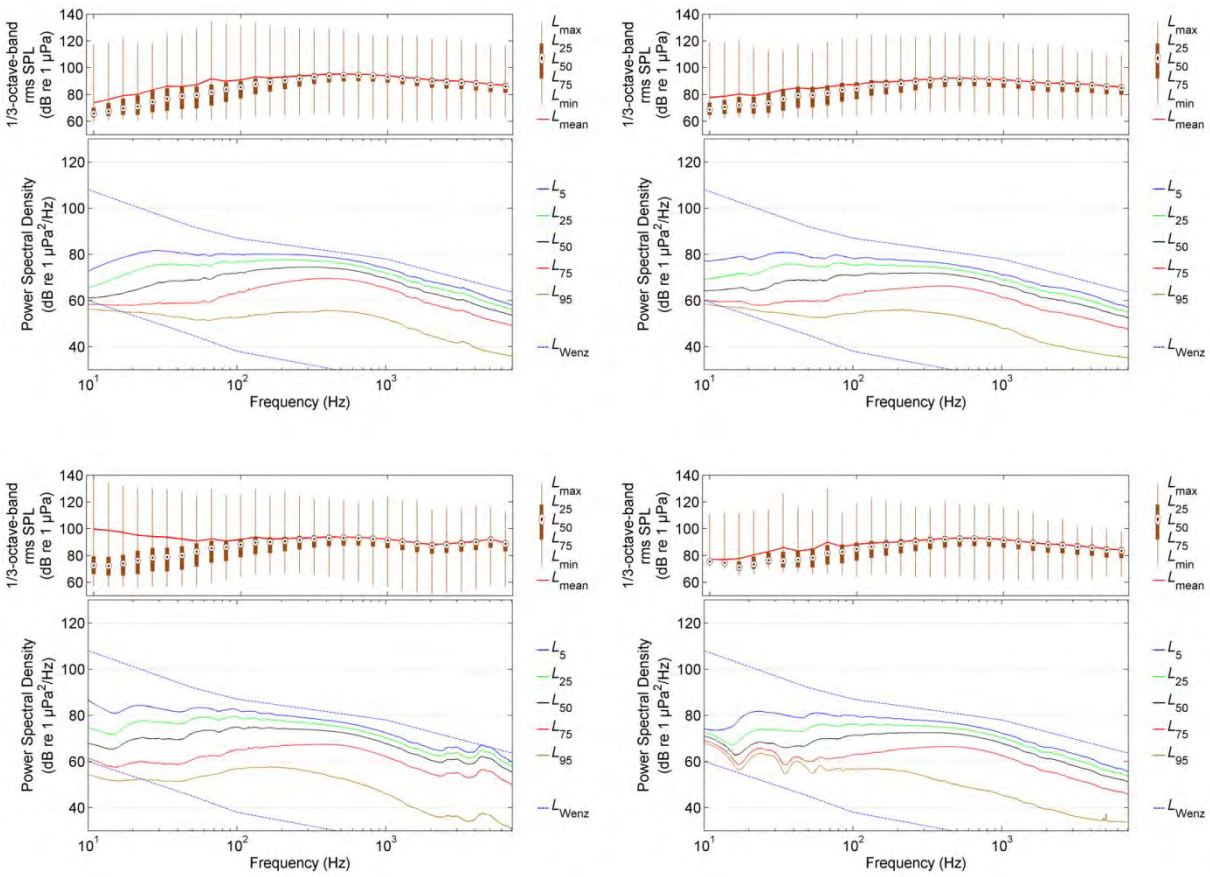

Figure B-10. 1/3-octave-band sound pressure levels and percentile 1 min power spectral density levels for summer 2014 stations. (Top left) KL01, (top right) PLN40, (bottom left) PLN60, and (bottom right) PLN80.

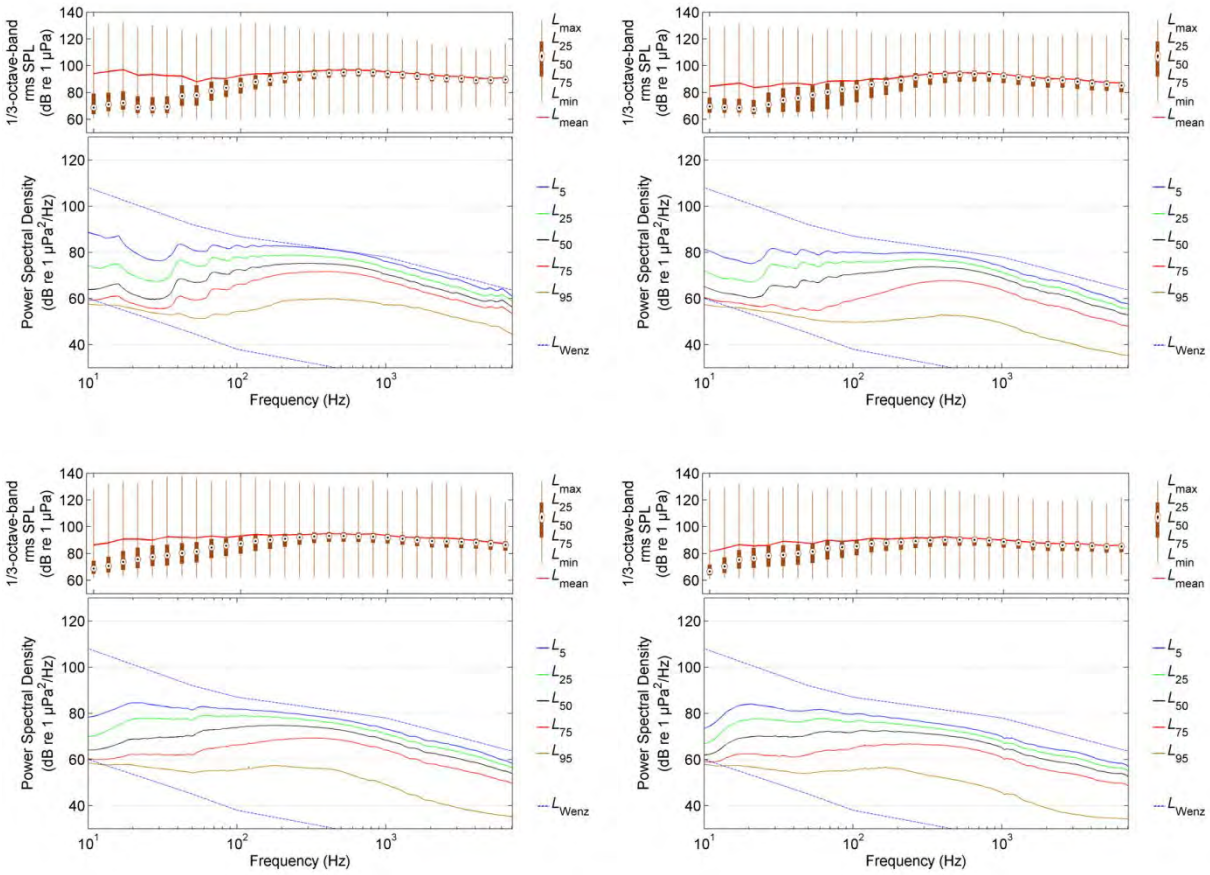

Figure B-11. 1/3-octave-band sound pressure levels and percentile 1 min power spectral density levels for summer 2014 stations. (Top left) CL5, (top right) CLN40, (bottom left) CLN90, and (bottom right) CLN120.

B.3.2. Broadband and Decade-Band Sound Pressure Levels and Spectrograms

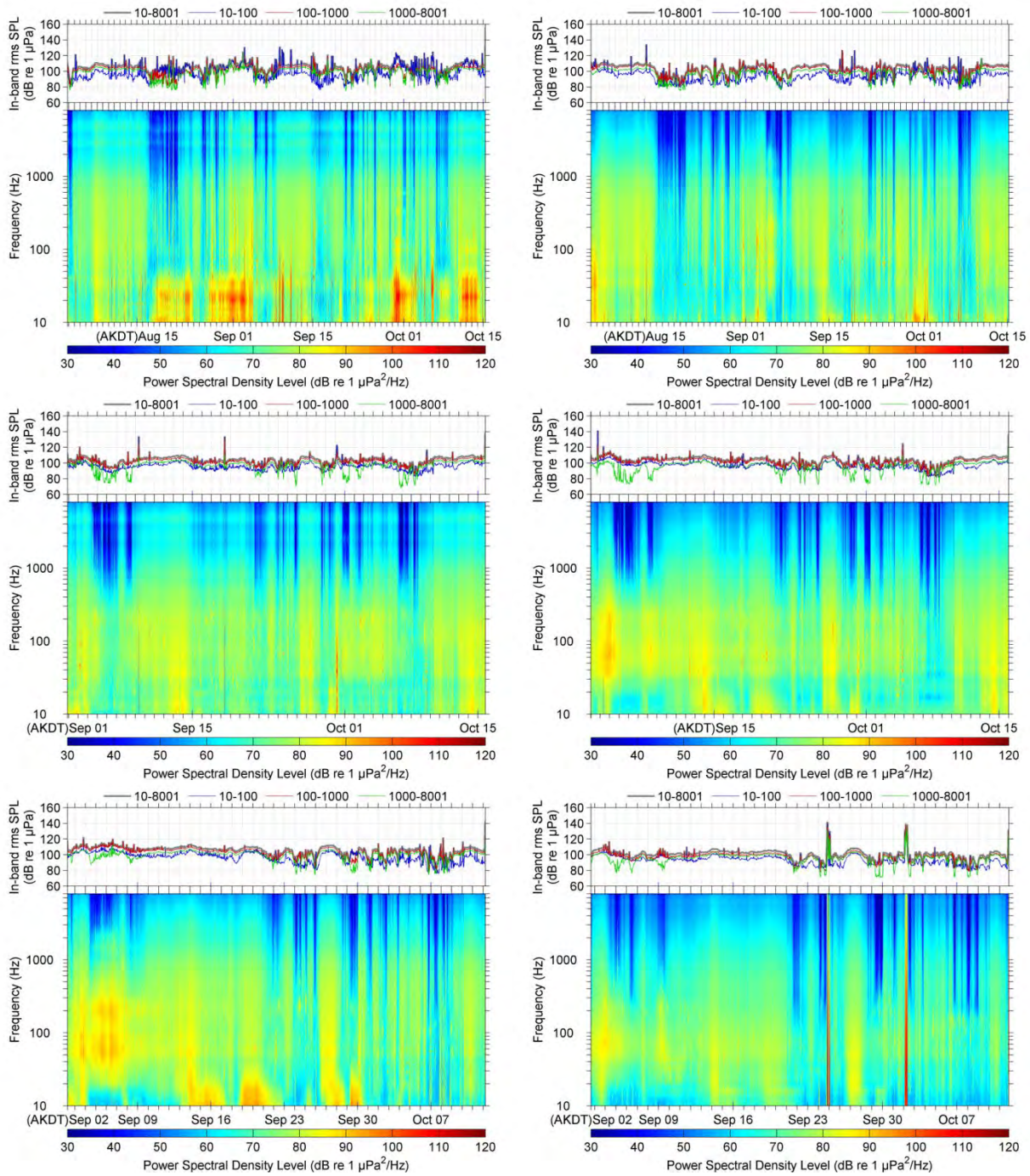

Figure B-12. Broadband and decade-band sound pressure levels (SPLs) and spectrograms for summer 2014 stations. (Top left) W10, (top right) W30, (middle left) W50, (middle right) WN20, (bottom left) WN40, and (bottom right) WN80.

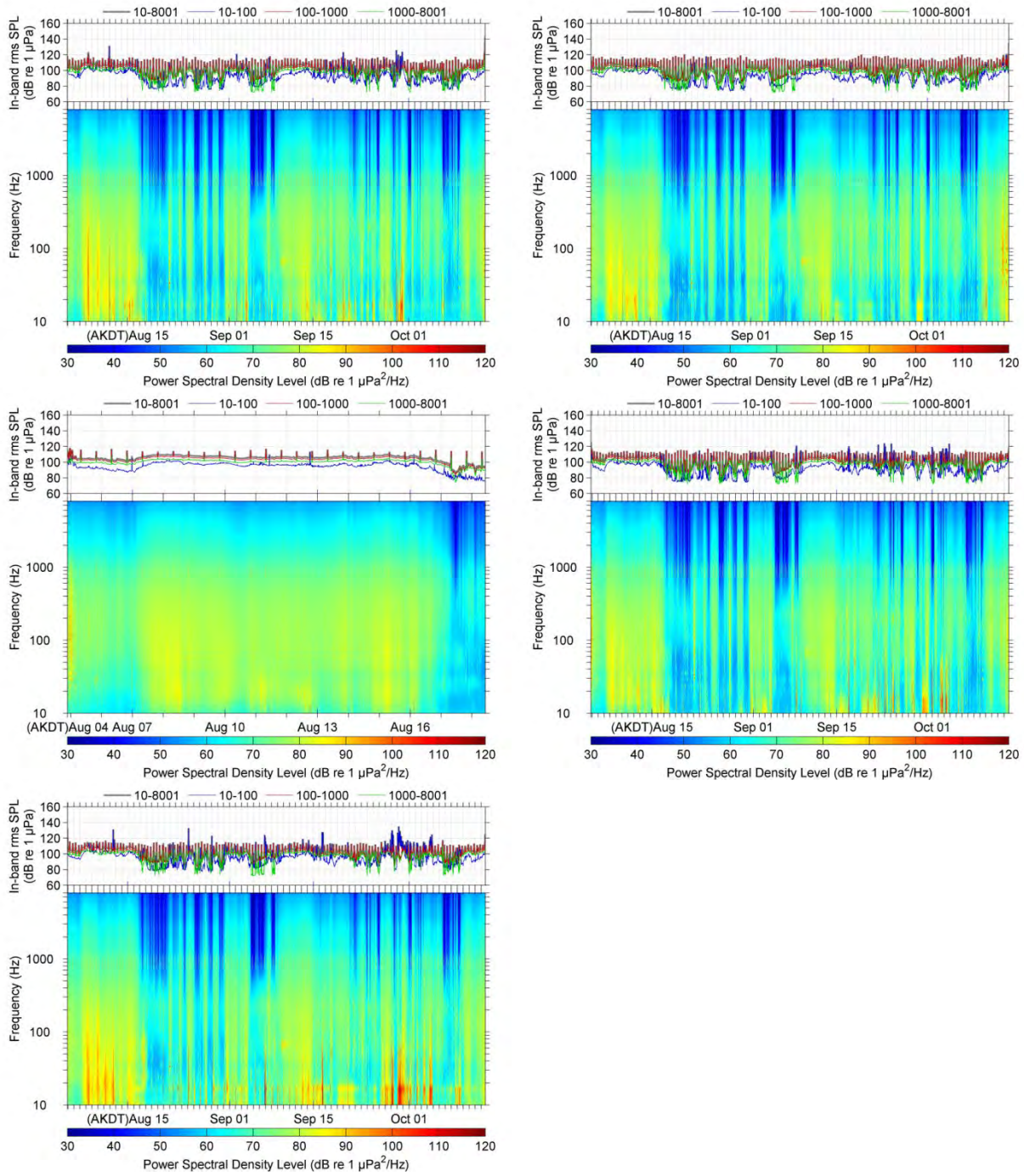

Figure B-13. Broadband and decade-band sound pressure levels (SPLs) and spectrograms for summer 2014 stations. (Top left) BGB, (top right) BGC, (middle left) BGD, (middle right) BGE, and (bottom) BGF.

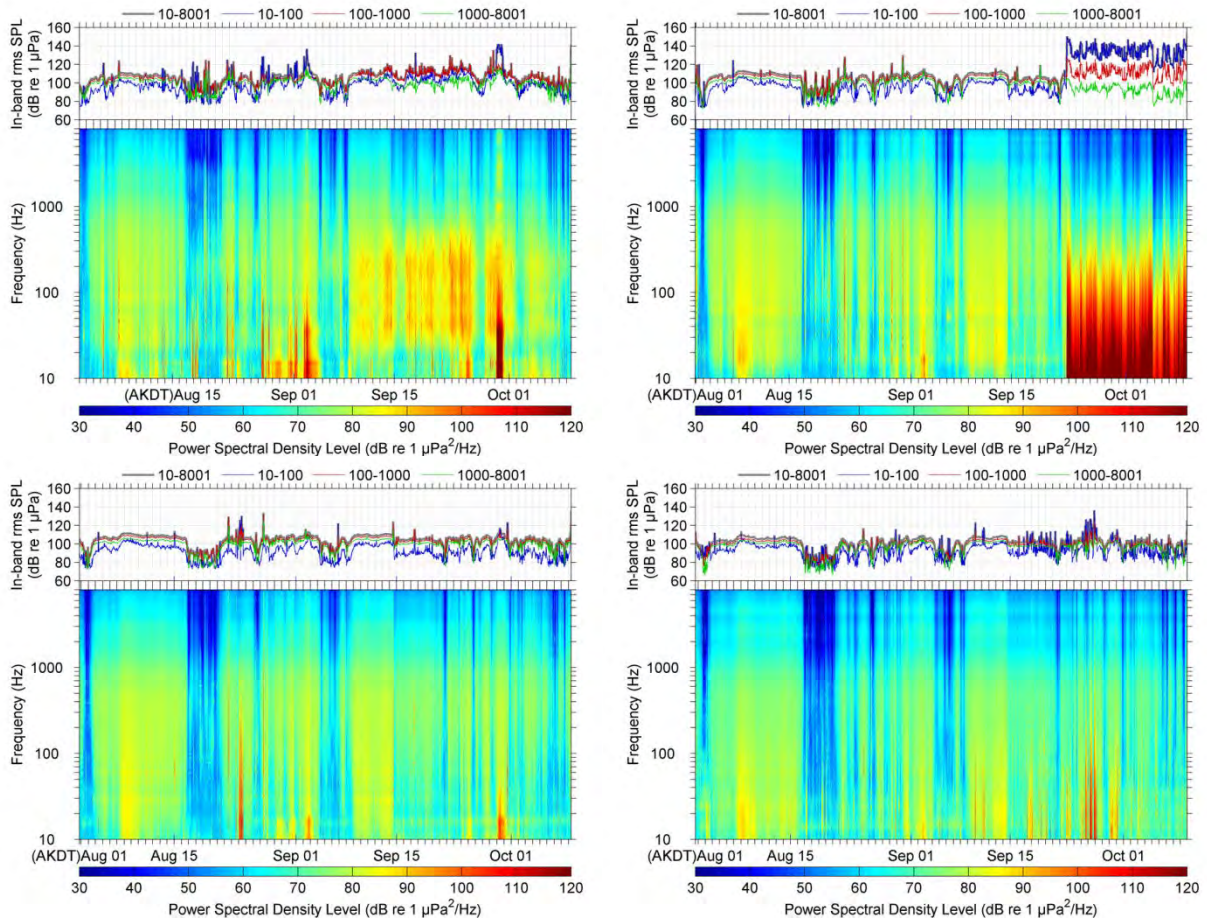

Figure B-14. Broadband and decade-band sound pressure levels (SPLs) and spectrograms for summer 2014 stations. (Top left) PL10, (top right) PL30, (bottom left) PL50, and (bottom right) PLN20.

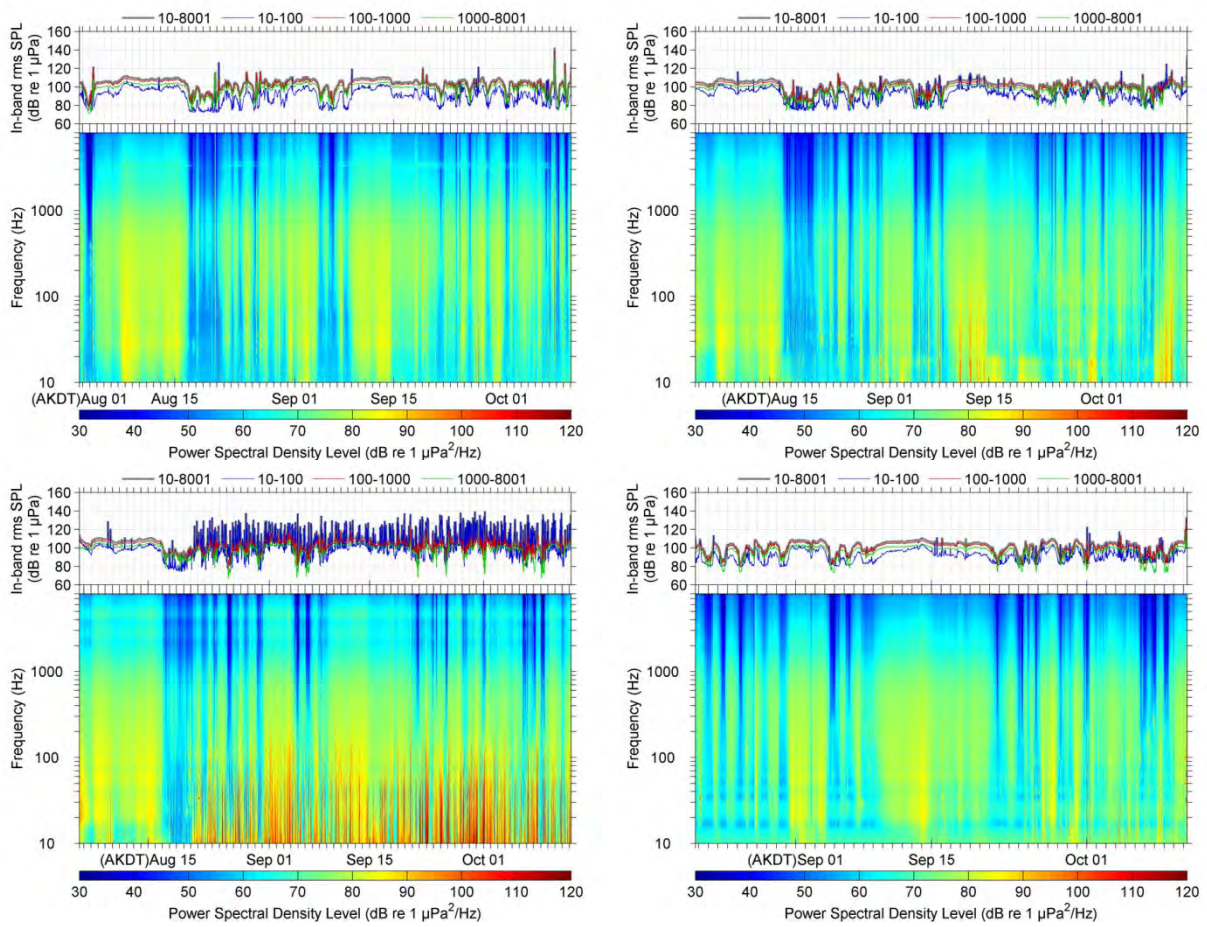

Figure B-15. Broadband and decade-band sound pressure levels (SPLs) and spectrograms for summer 2014 stations. (Top left) KL01, (top right) PLN40, (bottom left) PLN60, and (bottom right) PLN80.

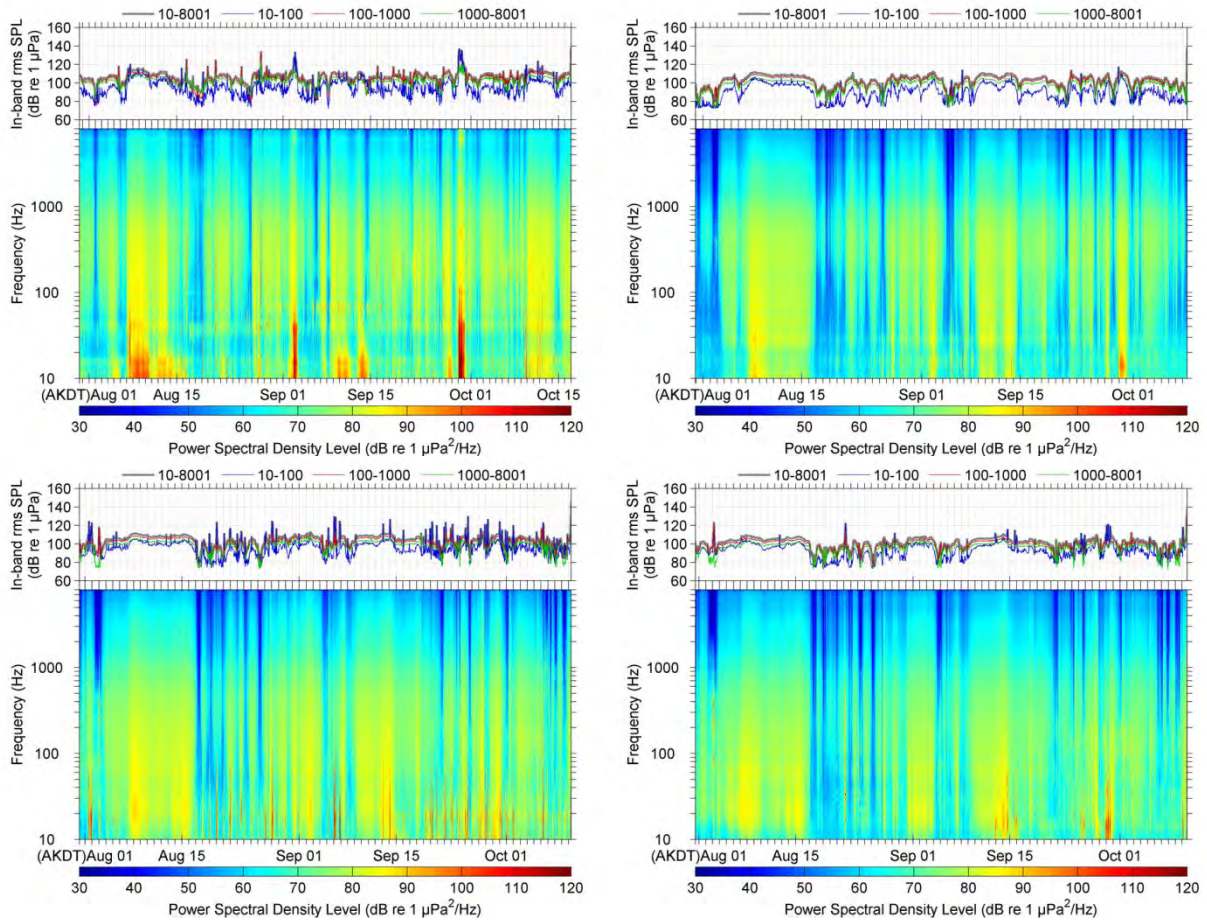

Figure B-16. Broadband and decade-band sound pressure levels (SPLs) and spectrograms for summer 2014 stations. (Top left) CL5, (top right) CLN40, (bottom left) CLN90, and (bottom right) CLN120.

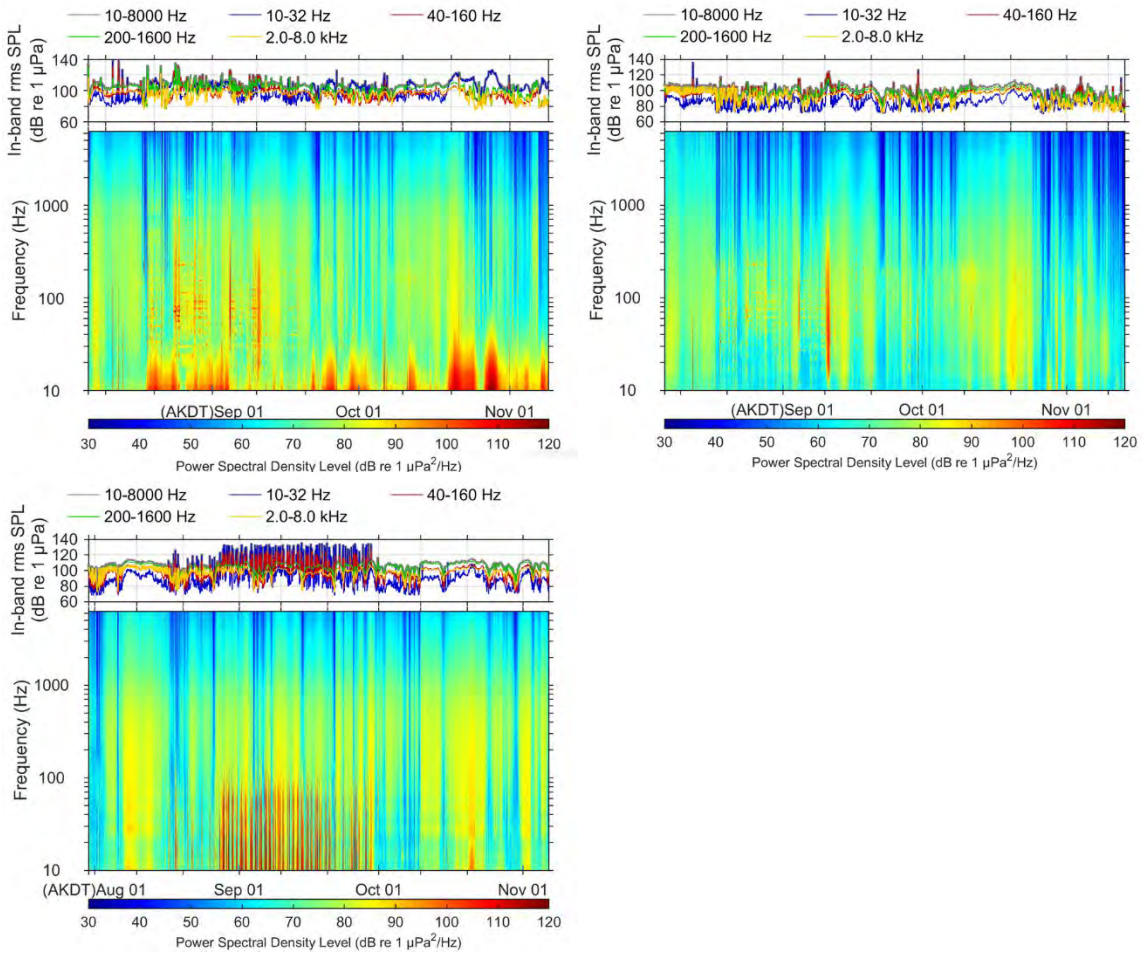

Figure B-17. Broadband and decade-band sound pressure levels (SPLs) and spectrograms for summer 2014 stations. (Top left) B5, (top right) B15, and (bottom) CL50.

B.3.3. Daily Cumulative Sound Exposure Level

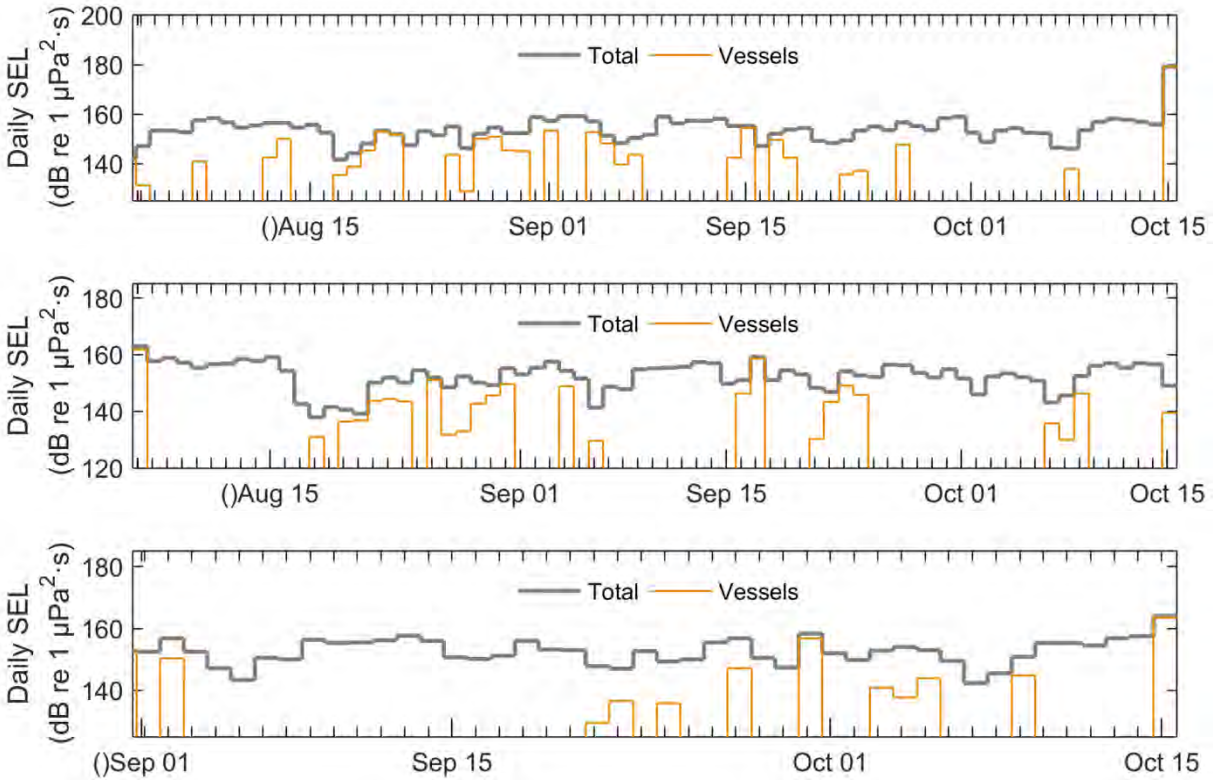

Figure B-18. Daily sound exposure levels (SEL 24 h) for summer 2014 stations. From top to bottom: W10, W30, and W50.

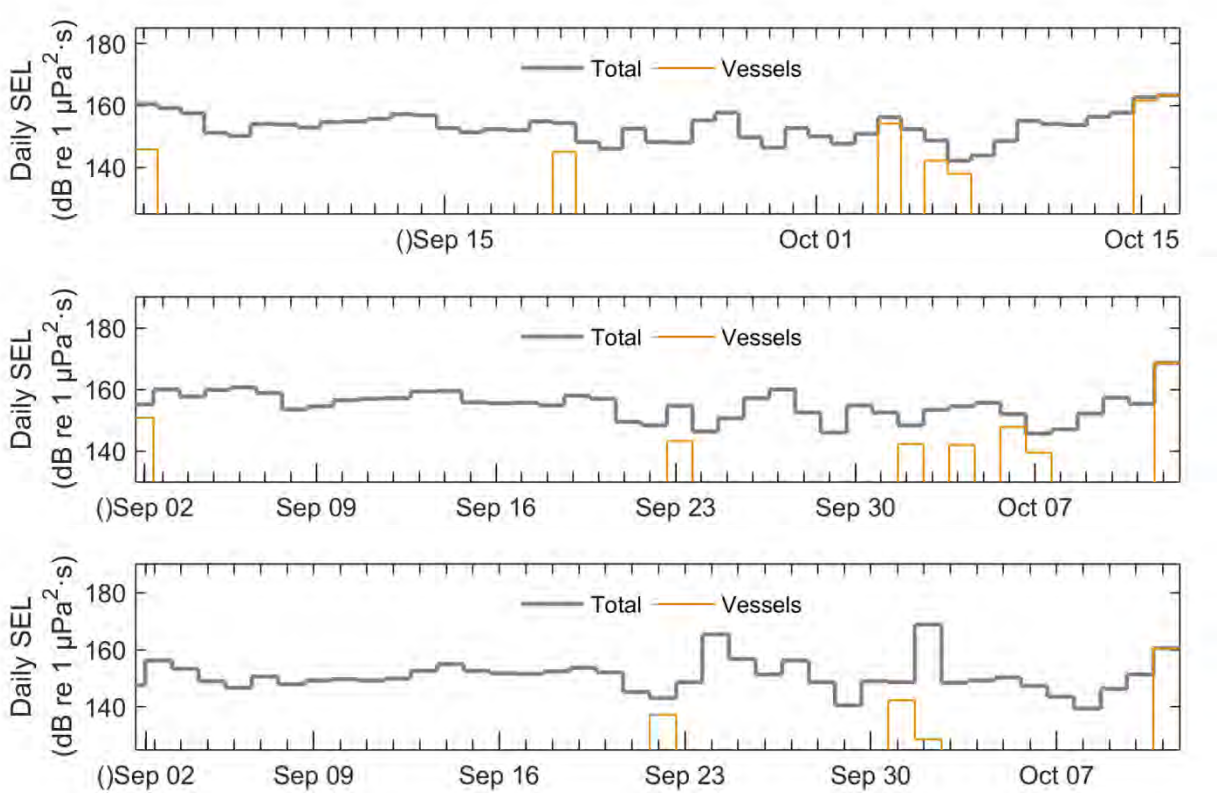

Figure B-19. Daily sound exposure levels (SEL 24 h) for summer 2014 stations. From top to bottom: WN20, WN40, and WN80.

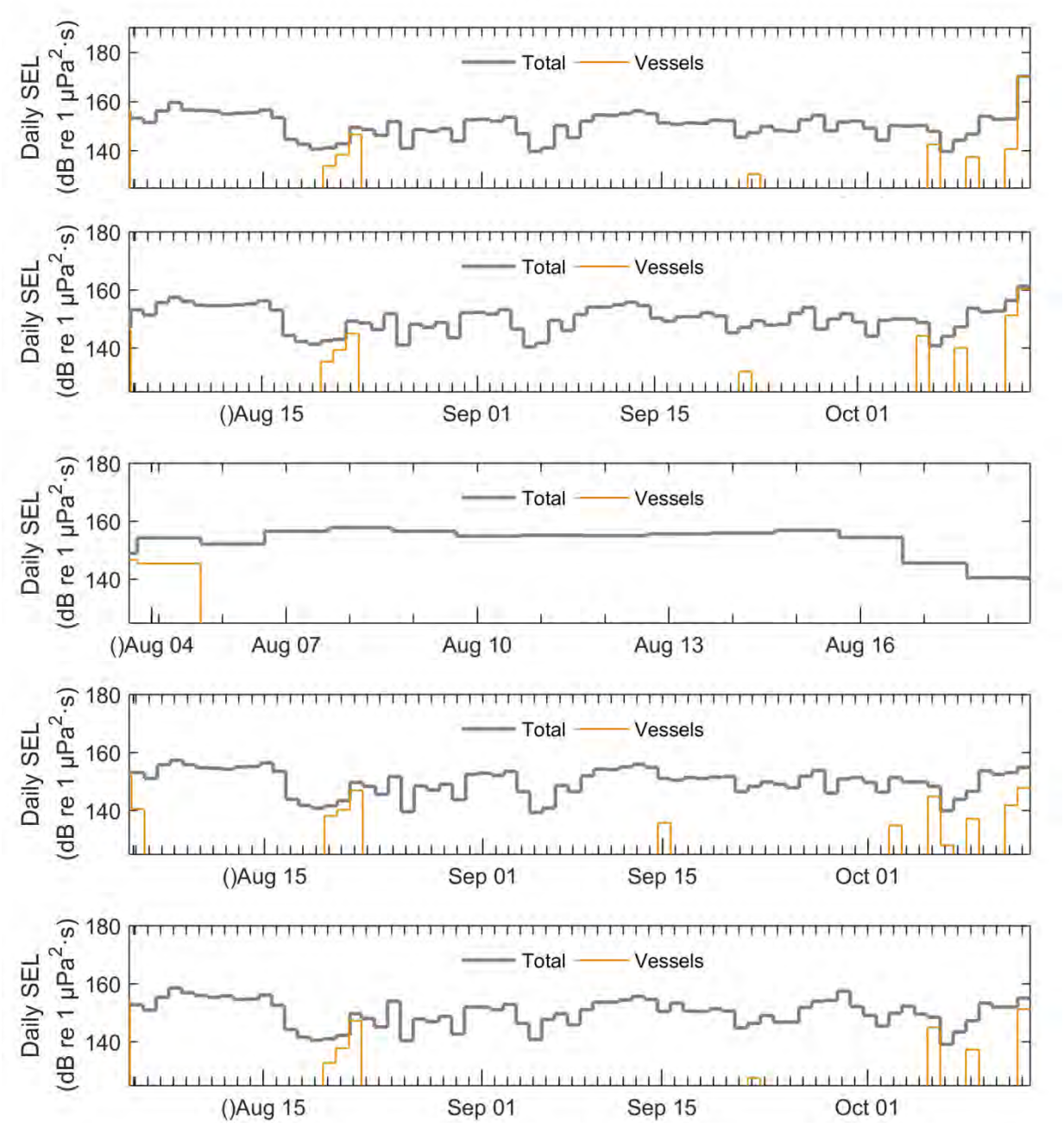

Figure B-20. Daily sound exposure levels (SEL 24 h) for summer 2014 stations. From top to bottom: BGB, BGC, BGD, BGE, and BGF.

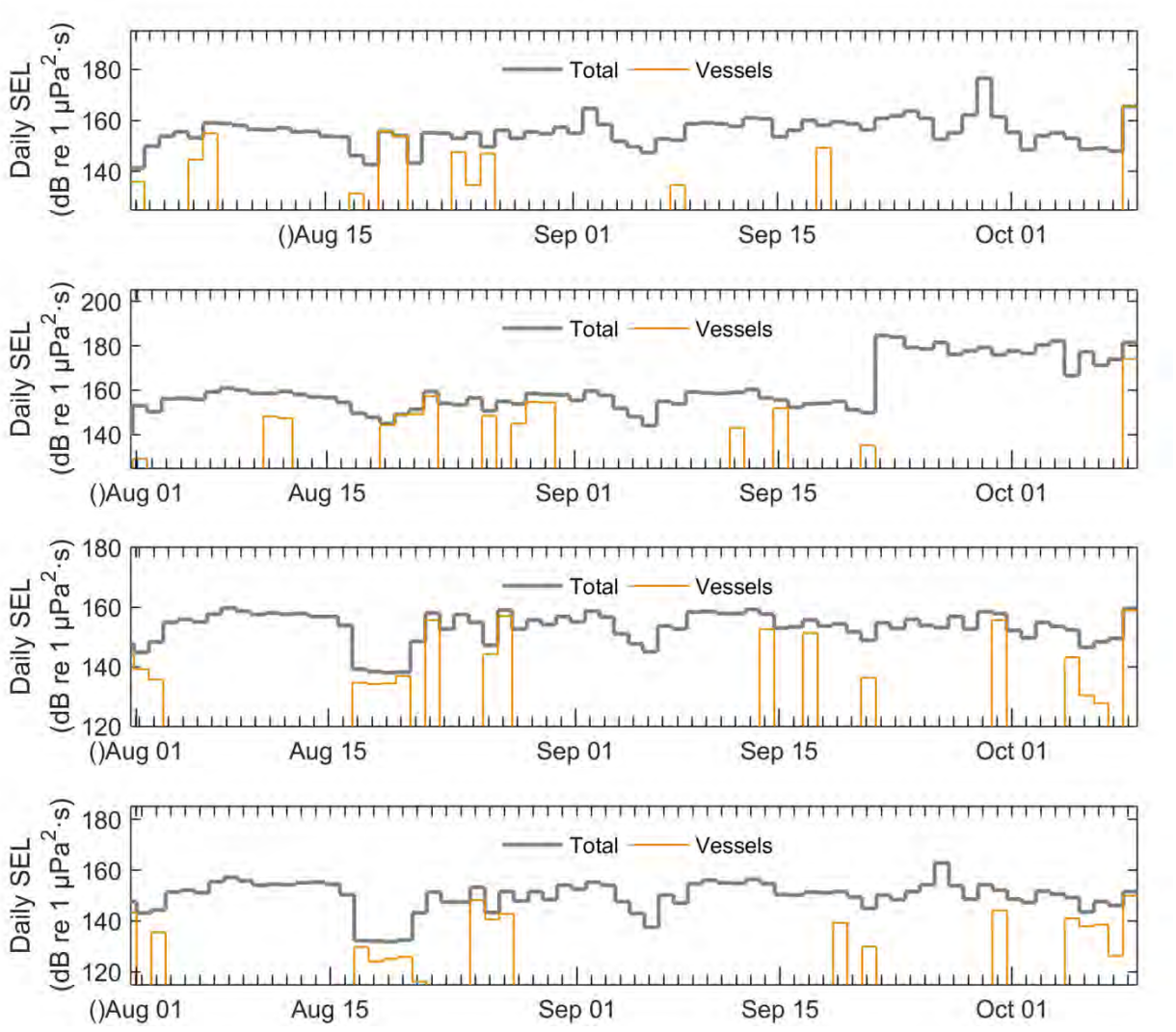

Figure B-21. Daily sound exposure levels (SEL 24 h) for summer 2014 stations. From top to bottom: PL10, PL30, PL50, and PLN20.

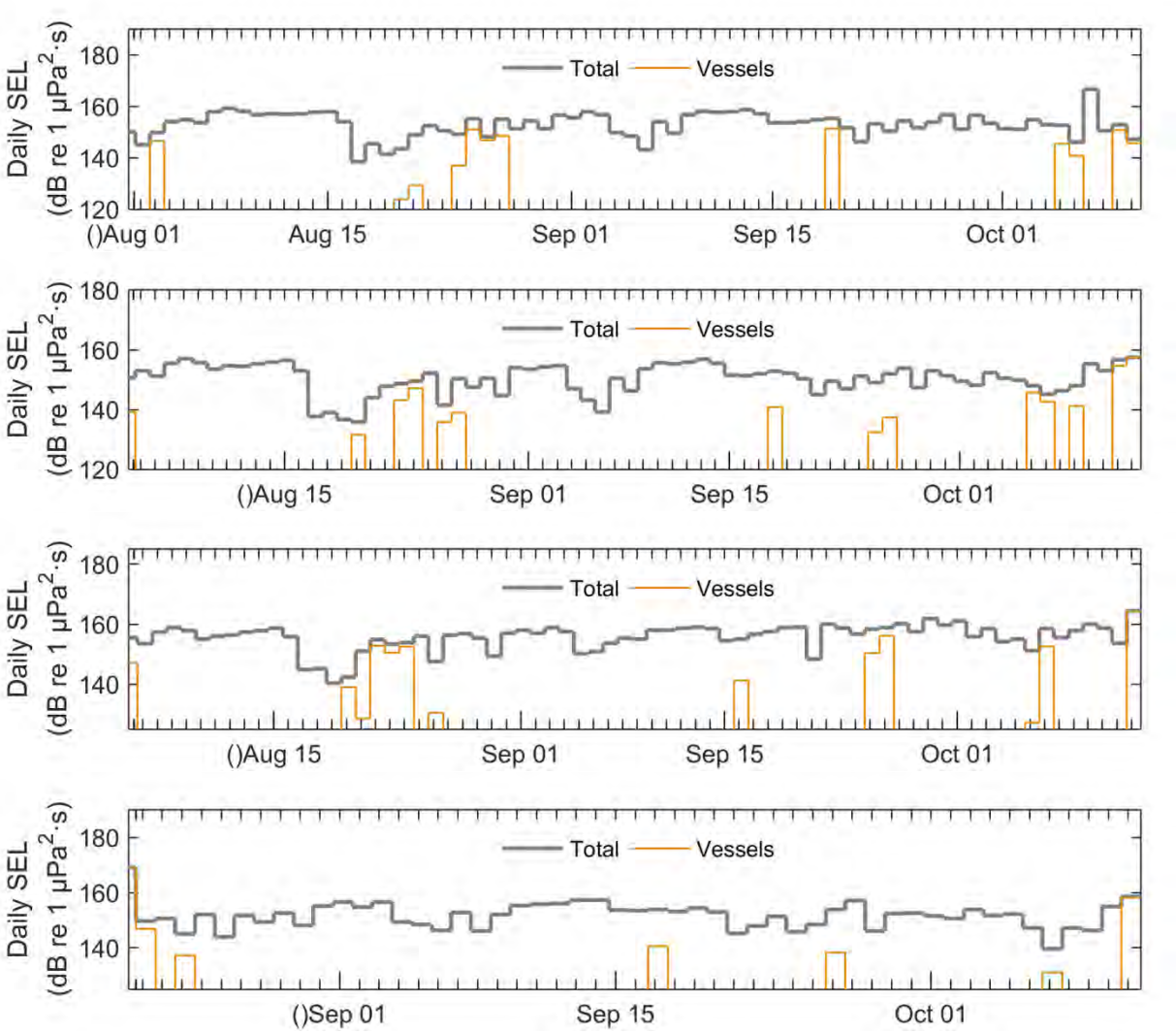

Figure B-22. Daily sound exposure levels (SEL 24 h) for summer 2014 stations. From top to bottom: KL01, PLN40, PLN60, and PLN80.

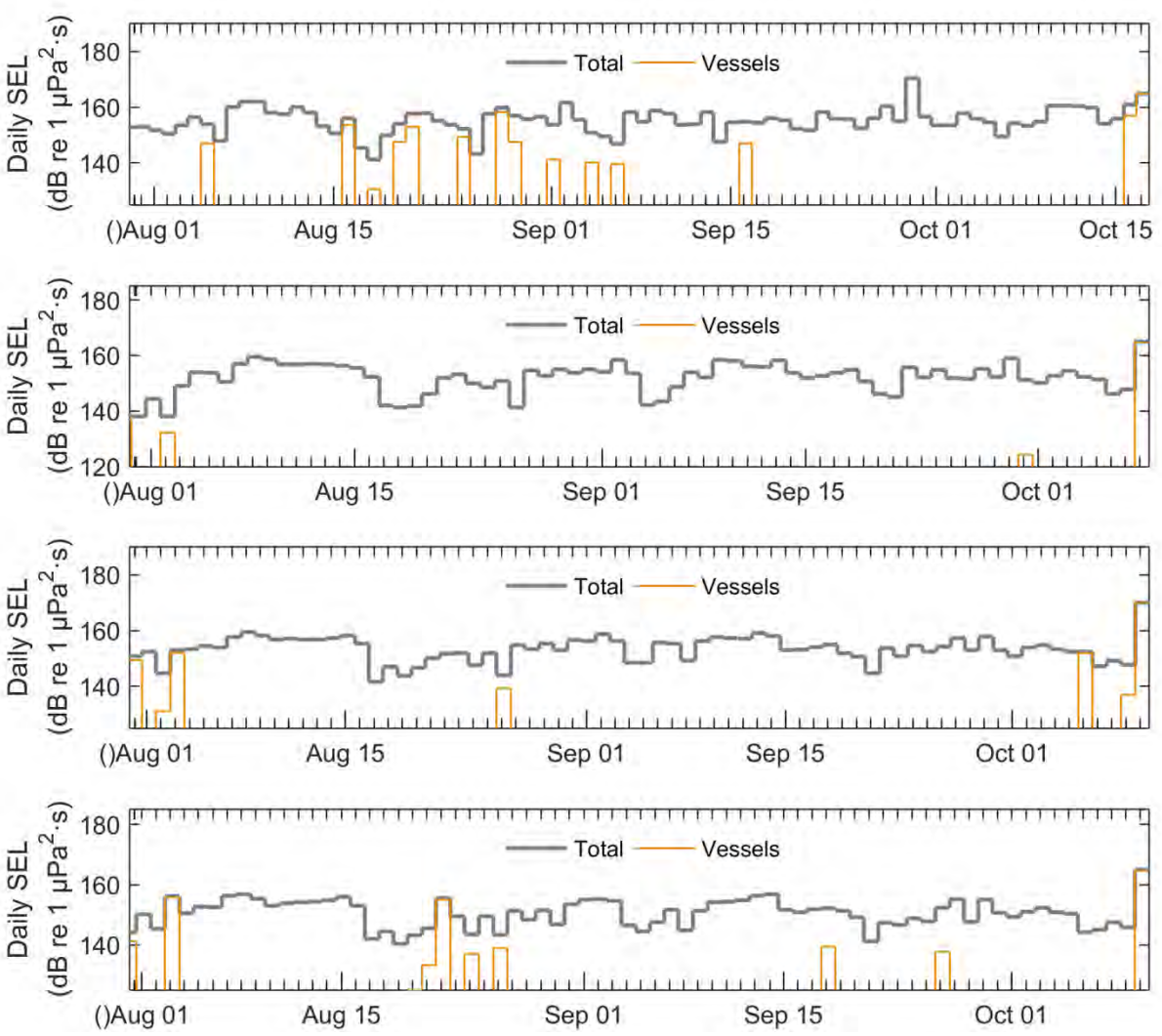

Figure B-23. Daily sound exposure levels (SEL 24 h) for summer 2014 stations. From top to bottom: CL5, CLN40, CLN90, and CLN120.

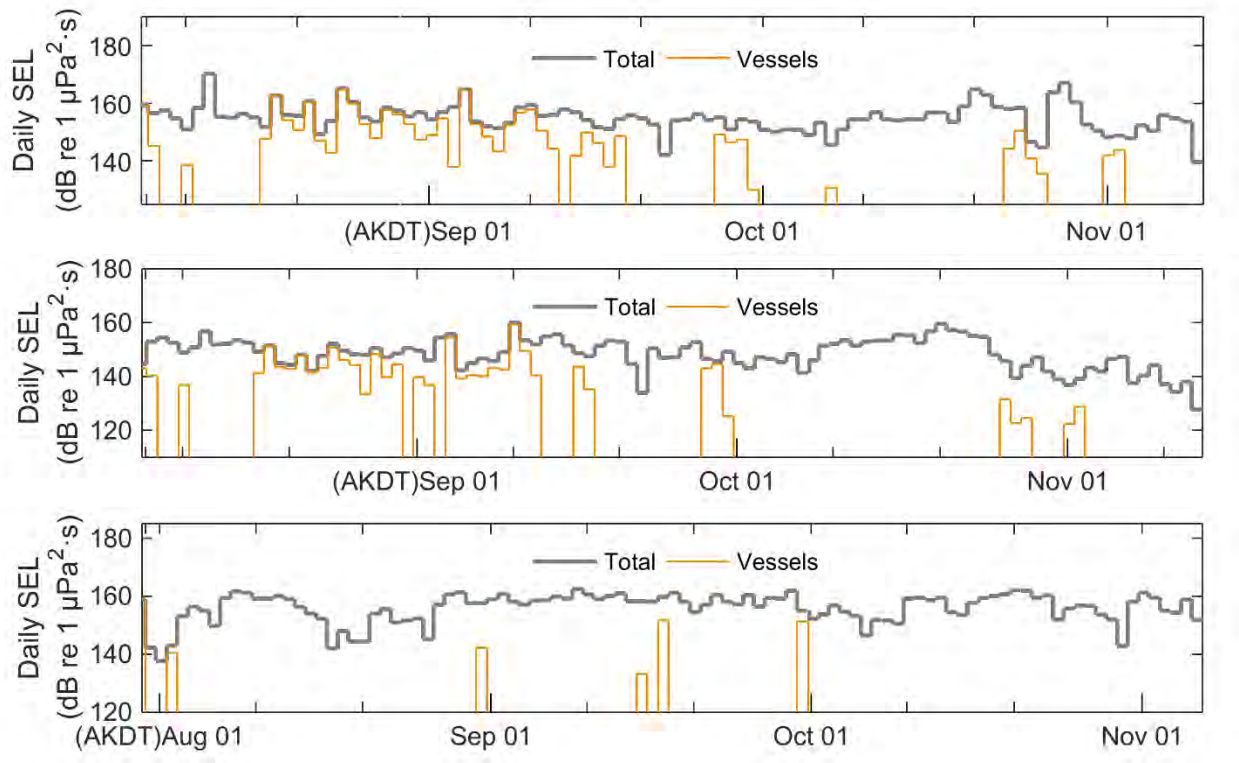

Figure B-24. Daily sound exposure levels (SEL 24 h) for summer 2014 stations. From top to bottom: B5, B15, and CL50.

B.3.4. Vessel Noise Detection

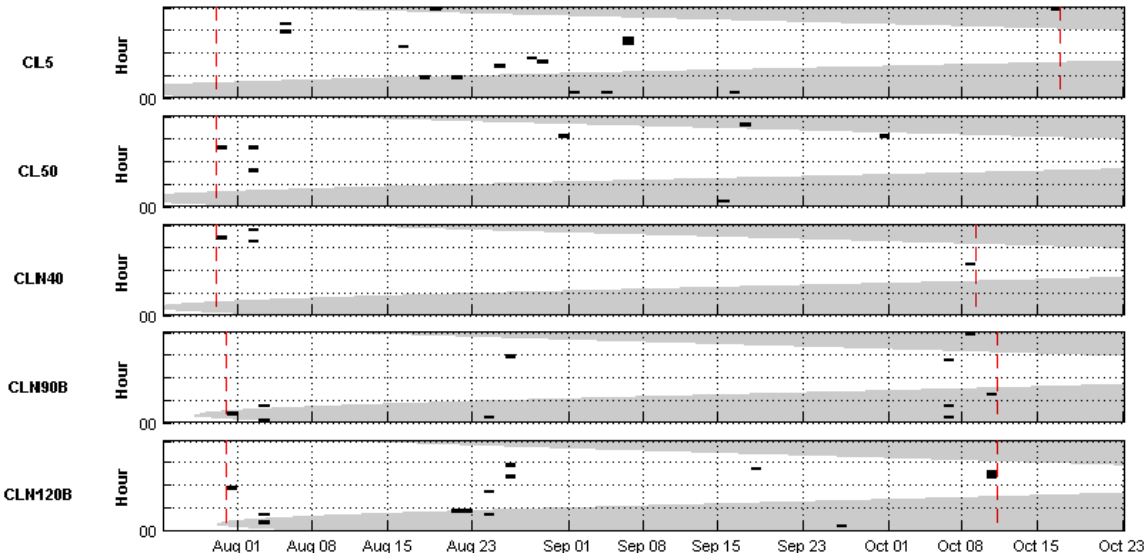

Figure B-25. Vessel detections each hour (vertical axis) versus date (horizontal axis) at five stations—CL5 to CLN120B—25 Jul to 23 Oct 2014. The grey areas indicate hours of darkness. Vertical dashed lines indicate AMAR deployment and retrieval dates.

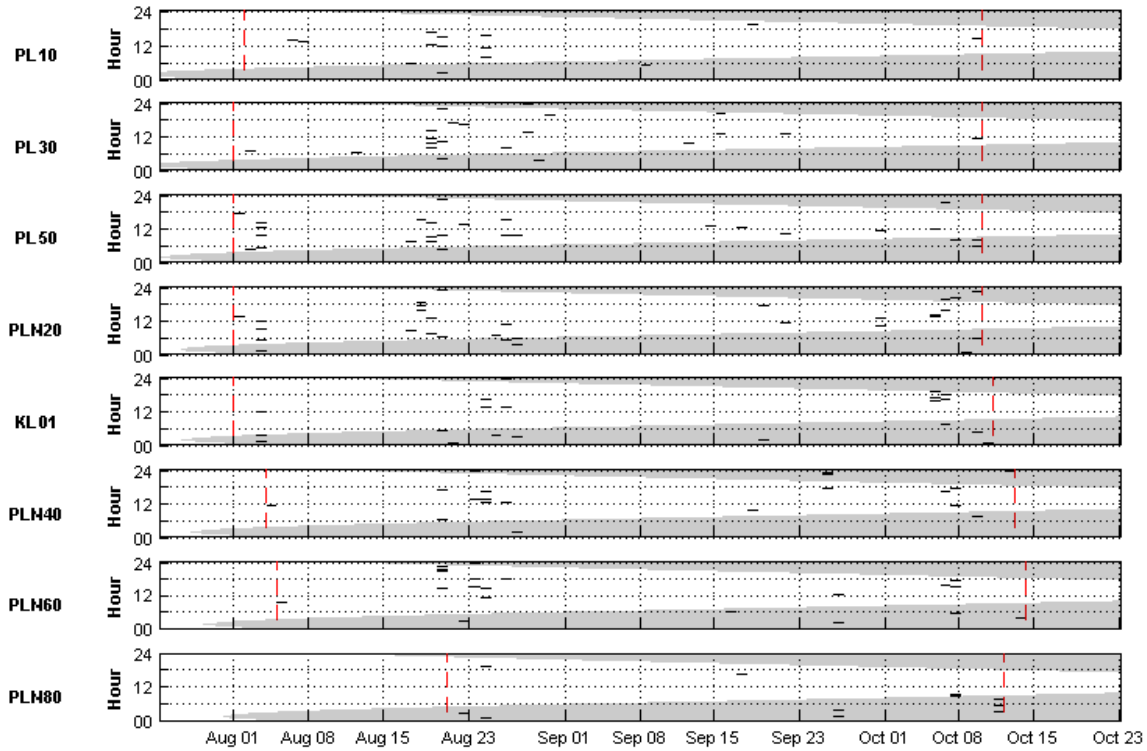

Figure B-26. Vessel detections each hour (vertical axis) versus date (horizontal axis) at eight stations—PL10 to PLN80—25 Jul to 23 Oct 2014. The grey areas indicate hours of darkness. Vertical dashed lines indicate AMAR deployment and retrieval dates.

Appendix C. Marine Mammal Detection Results

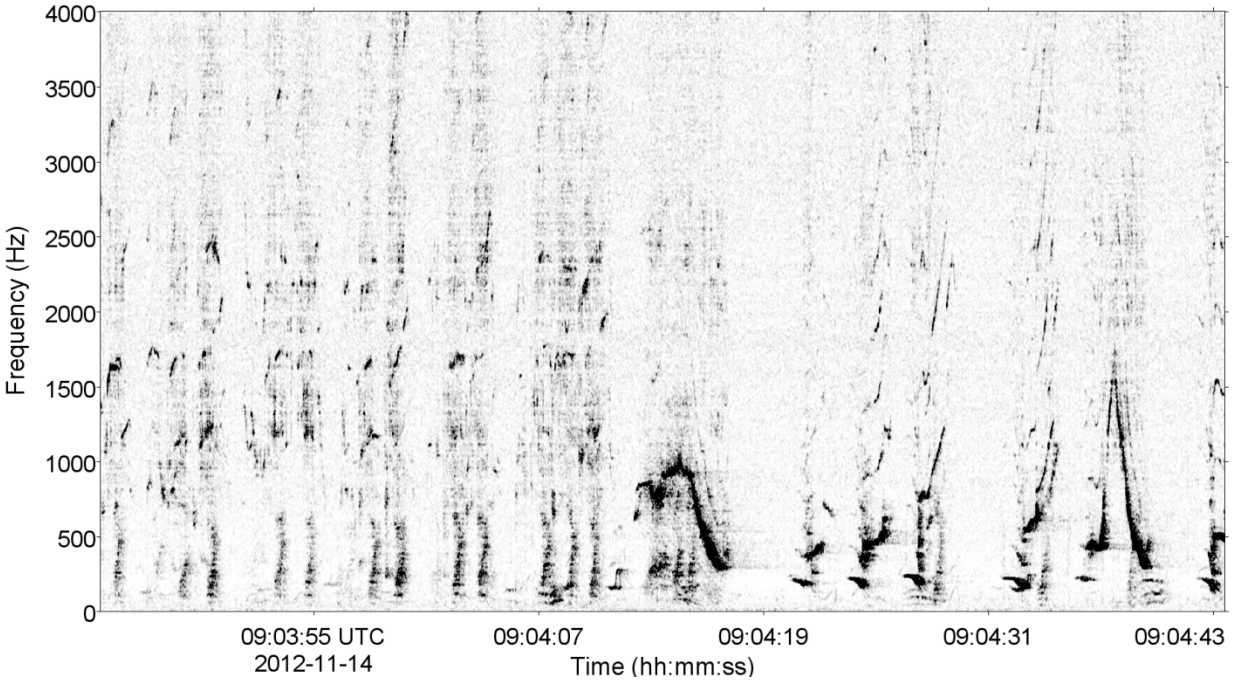

Figure C-1. Spectrogram of complex bowhead calls recorded at Station W35, 14 Nov 2012 (Frequency resolution: 2 Hz; Frame size: 0.128 s; Advance: 0.032 s; Hamming window).

Table C-1. Winter 2013–2014 bowhead call detections: Dates of first and last call detections, both possible (i.e., record start and end) and actual, and the number of days on which a call was detected manually for each recording station in the northeastern Chukchi Sea. The recorders operated for 5 min every 30 min.

Station	Record start	Fall 2013			Spring 2014			Record end
		First detection	Last detection	Detection days	First detection	Last detection	Detection days	
B05	11 Oct	11 Oct	12 Dec	47			0	26 Dec
WN40	12 Oct	16 Oct	22 Nov	34			0	18 Sept
W50	15 Oct	15 Oct	20 Dec	65	12 April	20 Sep	36	21 Sep
W10	11 Oct	12 Oct	16 Dec	50	13 Apr	15 Jul	65	2 Aug
PL50	19 Oct	19 Oct	28 Dec	56	4 Apr	15 Jul	43	12 Aug
PL10	18 Oct	22 Oct	4 Nov	5			0	27 Jan
CL05	20 Oct	31 Oct	4 Jan	54	9 Apr	18 May	27	11 Jun

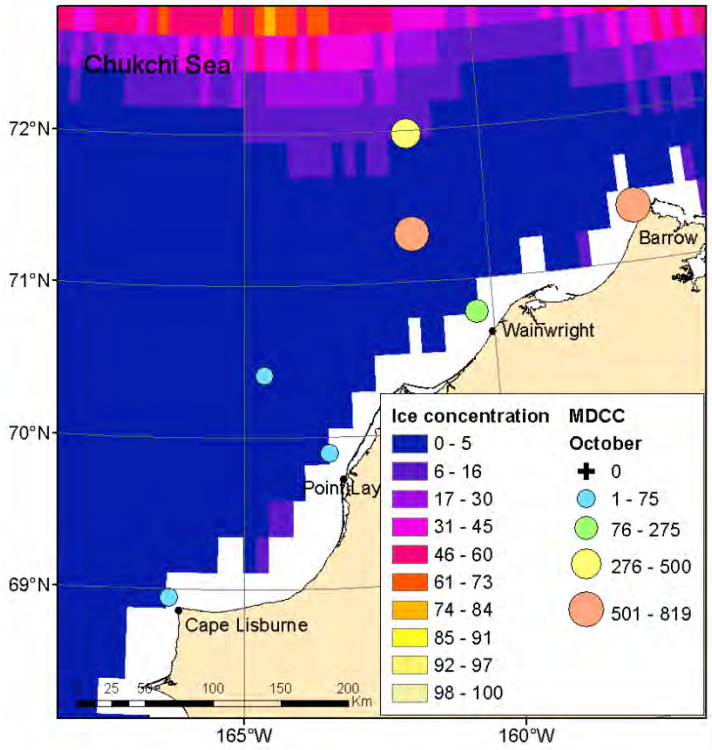

Figure C-2. Bowhead whale call count estimates* in the Chukchi Sea from 11-31 Oct 2013 at all winter 2013–2014 recorders. Ice data are for 23 Oct 2013.

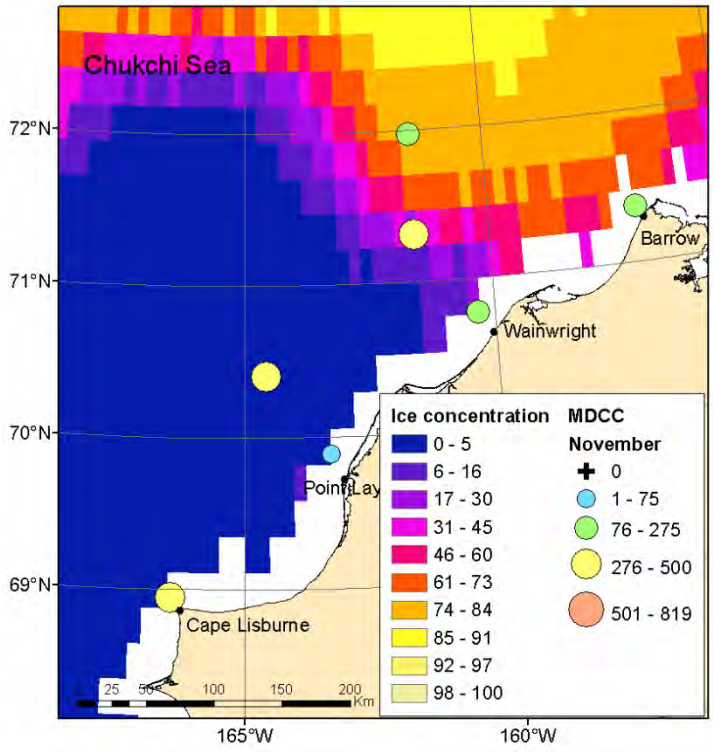

Figure C-3. Bowhead whale call count estimates* in the Chukchi Sea for November 2013 at all winter 2013–2014 recorders. Ice data are for 15 Nov 2013.

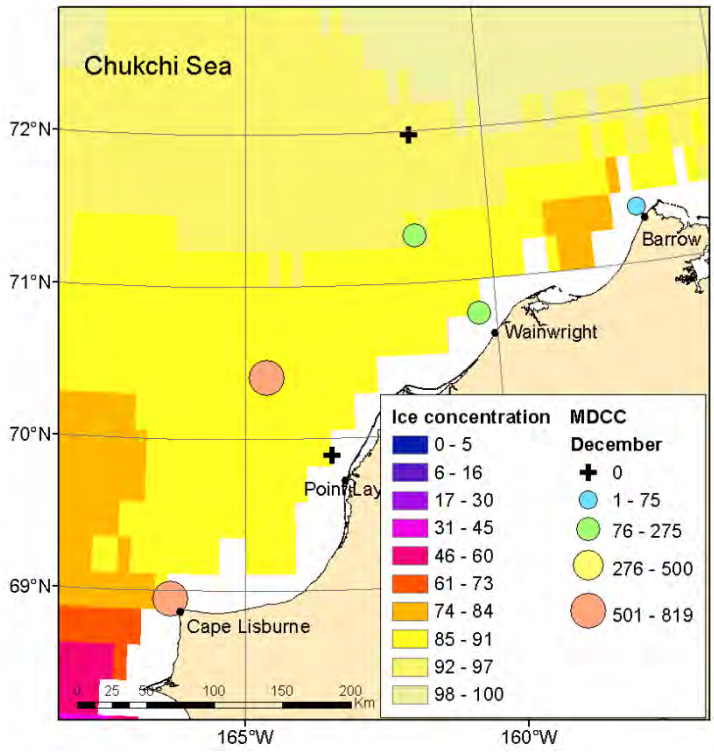

Figure C-4. Bowhead whale call count estimates* in the Chukchi Sea for December 2013 at all winter 2013–2014 recorders. Ice data are for 15 Dec 2013.

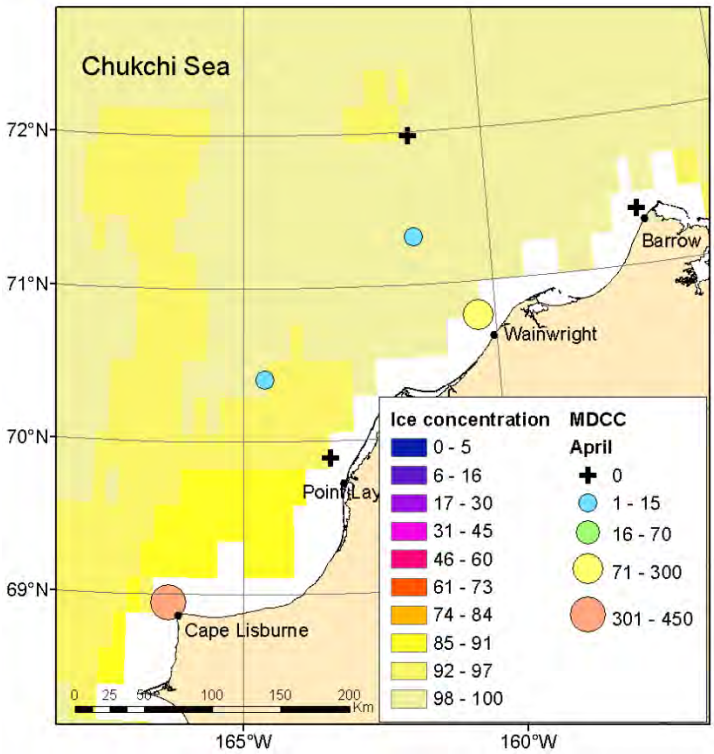

Figure C-5. Bowhead whale call count estimates* in the Chukchi Sea for April 2014 at all winter 2013–2014 recorders. Ice data are for 15 Apr 2014.

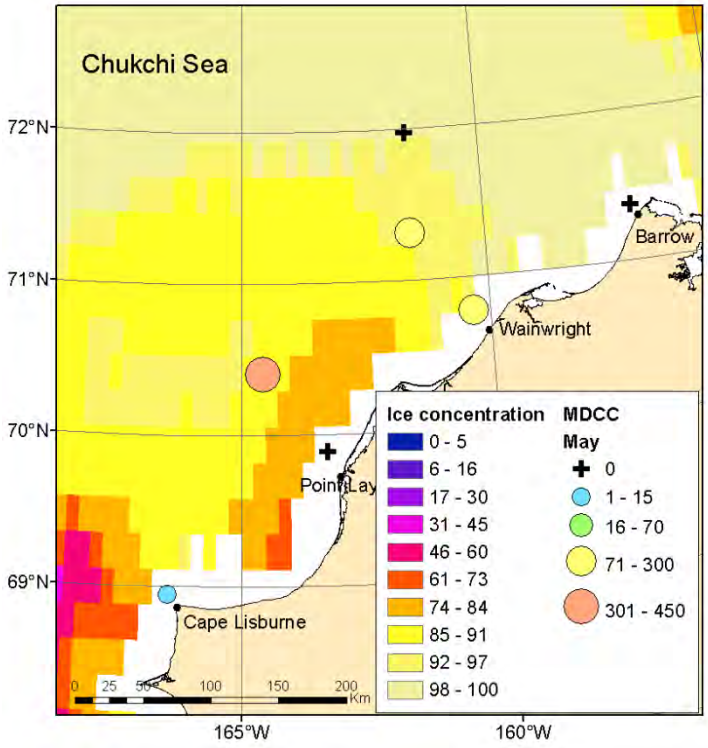

Figure C-6. Bowhead whale call count estimates* in the Chukchi Sea for May 2014 at all winter 2013–2014 recorders. Ice data are for 15 May 2014.

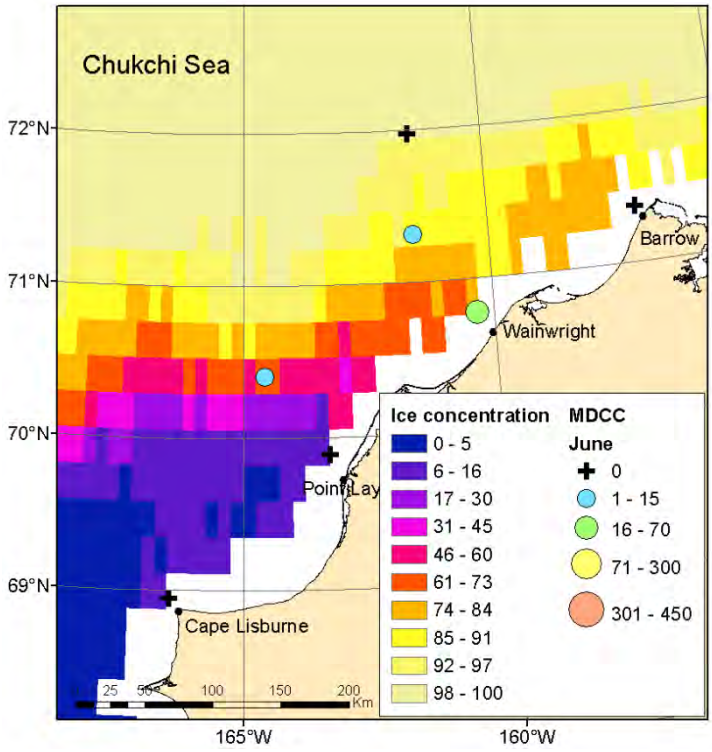

Figure C-7. Bowhead whale call count estimates* in the Chukchi Sea for June 2014 at all winter 2013–2014 recorders. Ice data are for 15 Jun 2014.

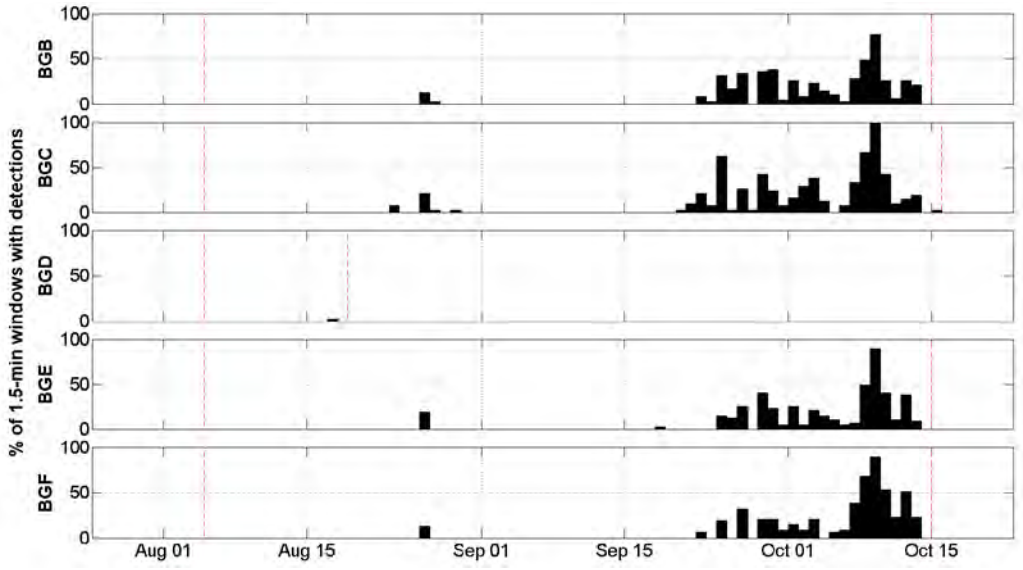

Figure C-8. Summer 2014 daily bowhead call detections in the Burger study area: Daily proportion of sound files with call detections based on the manual analysis of 5% of the acoustic data recorded late July through mid-October 2014. Forty-eight sound files were recorded daily. Vertical dashed lines indicate recording start and end. Stations without call detections were omitted.

Table C-2. Summer 2014 bowhead call detections: Dates of first and last call detections, both possible (i.e., record start and end) and actual, and the percent of days on which a call was detected for each recording station in the northeastern Chukchi Sea. Stations without call detections are omitted.

Station	Record start	First detection	Last detection	Record end	Detection days	% Days with detection
BGB	5 Aug	26 Aug	13 Oct	14 Oct	23	34.0
BGC	5 Aug	23 Aug	15 Oct	15 Oct	28	39.5
BGD	5 Aug	17 Aug	17 Aug	18 Aug	1	7.4
BGE	5 Aug	26 Aug	13 Oct	14 Oct	21	30.0
BGF	5 Aug	26 Aug	13 Oct	14 Oct	19	27.2
B15	6 Aug	7 Aug	13 Nov	13 Nov	71	72.4
B5	6 Aug	9 Aug	5 Nov	9 Nov	63	68.1
WN80	2 Sep	21 Sep	12 Oct	12 Oct	20	48.8
WN40	2 Sep	12 Sep	12 Oct	12 Oct	20	48.8
WN20	2 Sep	9 Sep	16 Oct	16 Oct	23	52.3
W50	2 Sep	17 Sep	16 Oct	16 Oct	22	50.1
W30	6 Aug	17 Aug	15 Oct	15 Oct	27	38.5
W10	3 Aug	12 Aug	15 Oct	15 Oct	19	25.8
PLN80	21 Aug	27 Aug	9 Oct	11 Oct	17	33.5
PLN60	5 Aug	24 Aug	13 Oct	13 Oct	26	37.7
PLN40	4 Aug	14 Aug	12 Oct	13 Oct	27	38.8
KL01	1 Aug	21 Sep	9 Oct	10 Oct	19	27.3
PLN20	1 Aug	24 Aug	10 Oct	10 Oct	21	30.3
PL50	2 Aug	22 Sep	8 Oct	9 Oct	9	13.1
CLN120B	31 Jul	19 Aug	10 Oct	11 Oct	23	32.3
CLN90B	31 Jul	21 Sep	10 Oct	10 Oct	19	26.7
CLN40	31 Jul	28 Sep	8 Oct	8 Oct	6	8.6
CL50	31 Jul	13 Oct	6 Nov	6 Nov	9	9.3

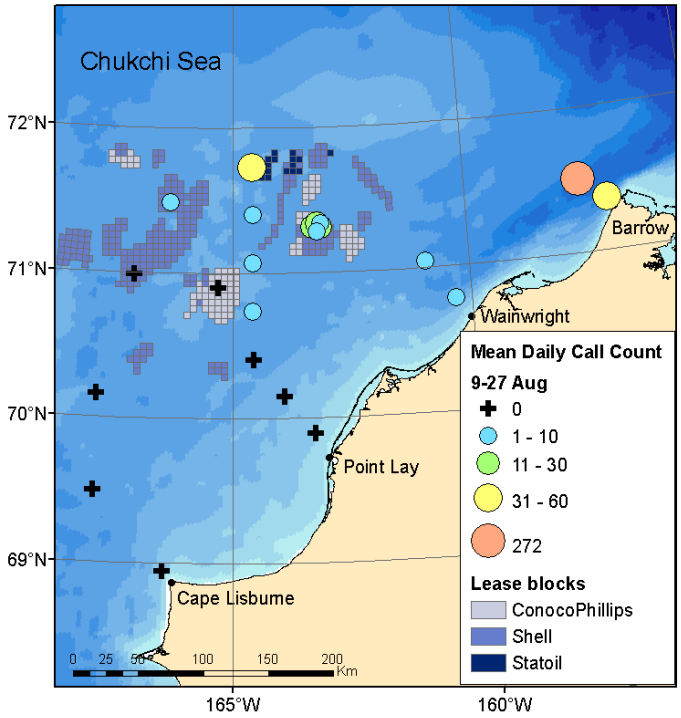

Figure C-9. Mean daily bowhead whale call counts (calculated as the sum of automated call detections in all files with manual detections divided by the number of active recording days) from 9-27 Aug at all summer 2014 stations in the northeastern Chukchi Sea.

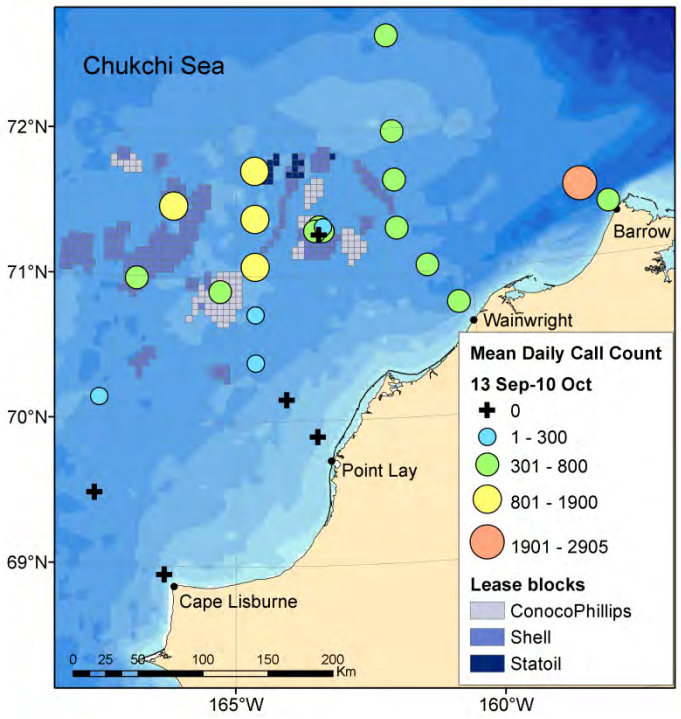

Figure C-10. Mean daily bowhead whale call counts (calculated as the sum of automated call detections in all files with manual detections divided by the number of active recording days) from 13 Sep to 10 Oct at all summer 2014 stations in the northeastern Chukchi Sea.

Table C-3. Winter 2013–2014 walrus call detections: Dates of first and last call detections, both possible (i.e., record start and end) and actual, and the number of days on which a call was detected manually for each recording station in the northeastern Chukchi Sea. The recorders operated for 5 min every 30 min.

Station	Record start	Fall 2013			Spring 2014			Record end
		First detection	Last detection	Detection days	First detection	Last detection	Detection days	
WN40	12 Oct	16 Oct	15 Nov	3	16 Jun	18 Sep	92	18 Sep
W50	15 Oct	--	--	0	5 Jun	20 Sep	100	21 Sep
W10	11 Oct	22 Oct	22 Oct	1	6 Jun	31 Jul	39	2 Aug
PL50	19 Oct	22 Oct	22 Nov	4	11 Jun	26 Jul	32	2 Aug
PL10	18 Oct	21 Oct	24 Oct	4	--	--	0	27 Jan
CL05	20 Oct	21 Oct	23 Oct	3	9 Jun	9 Jun	1	11 Jun

-- No detections

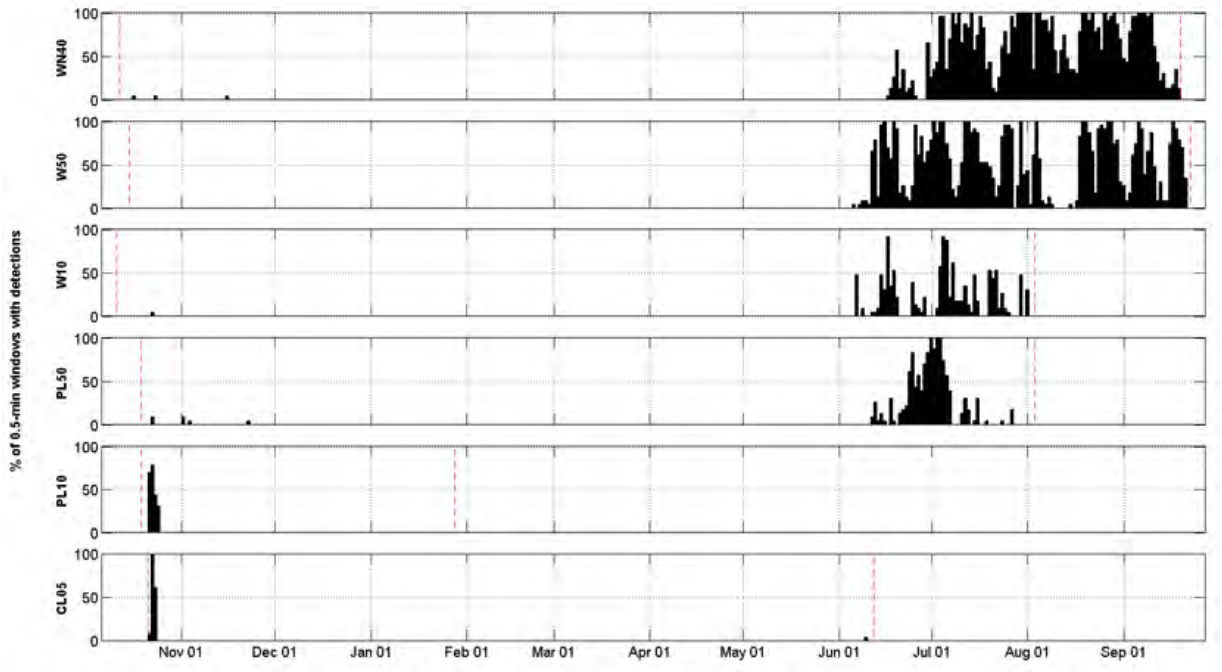

Figure C-11. Winter 2013–2014 daily walrus call detections: Daily proportion of sound files with call detections based on the manual analysis of 5% of the acoustic data recorded October 2013 through September 2014 in the northeastern Chukchi Sea for each station. Forty-eight sound files were recorded daily. Vertical dashed lines indicate recording start and end. Stations without call detections were omitted.

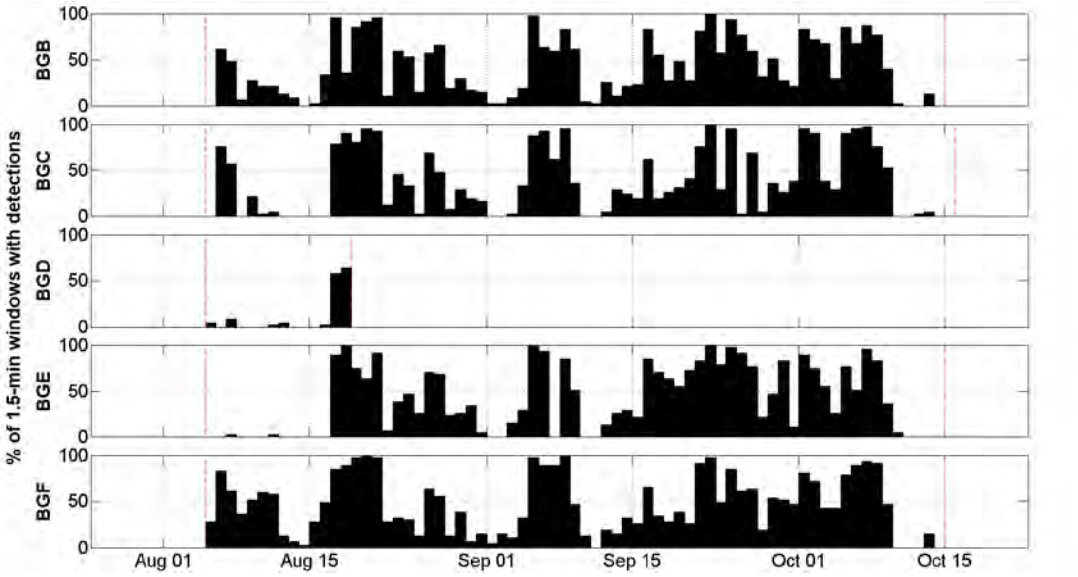

Figure C-12. Summer 2014 daily walrus call detections in the Burger study area: Daily proportion of sound files with call detections based on the manual analysis of 5% of the acoustic data recorded late July through mid-October 2014. Forty-eight sound files were recorded daily. Vertical dashed lines indicate recording start and end. Stations without call detections were omitted.

Table C-4. Summer 2014 walrus call detections: Dates of first and last call detections, both possible (i.e., record start and end) and actual, and the percent of days on which a call was detected for each recording station in the northeastern Chukchi Sea. Stations without call detections are omitted.

Station	Record start	First detection	Last detection	Record end	Detection days	% Days with detection
BGB	5 Aug	6 Aug	13 Oct	14 Oct	66	94.6
BGC	5 Aug	6 Aug	13 Oct	15 Oct	57	80.4
BGD	5 Aug	5 Aug	18 Aug	18 Oct	7	51.5
BGE	5 Aug	7 Aug	10 Oct	14 Oct	52	74.3
BGF	5 Aug	5 Aug	13 Oct	14 Oct	66	94.6
B15	6 Aug	7 Aug	2 Nov	13 Nov	20	20.4
B5	6 Aug	11 Aug	4 Nov	9 Nov	13	13.8
WN80	2 Sep	3 Sep	12 Oct	12 Oct	40	97.6
WN40	2 Sep	2 Sep	11 Oct	12 Oct	40	97.6
WN20	2 Sep	2 Sep	13 Oct	16 Oct	40	90.9
W50	2 Sep	2 Sep	9 Oct	16 Oct	38	86.6
W30	6 Aug	15 Aug	8 Oct	15 Oct	46	65.5
W10	3 Aug	3 Aug	13 Oct	15 Oct	40	54.3
PLN80	21 Aug	23 Aug	10 Oct	11 Oct	39	76.9
PLN60	5 Aug	5 Aug	10 Oct	13 Oct	37	53.7
PLN40	4 Aug	17 Aug	9 Oct	13 Oct	38	54.7
KL01	1 Aug	2 Aug	8 Oct	10 Oct	21	30.2
PLN20	1 Aug	2 Aug	8 Oct	10 Oct	40	57.6
PL50	2 Aug	2 Aug	7 Oct	9 Oct	25	36.4
PL30	2 Aug	2 Aug	23 Sep	9 Oct	37	54
PL10	2 Aug	2 Aug	9 Oct	9 Oct	62	90.4
CLN120B	31 Jul	1 Aug	10 Oct	11 Oct	51	71.6
CLN90B	31 Jul	31 Jul	10 Oct	10 Oct	49	68.8
CLN40	31 Jul	31 Jul	8 Oct	8 Oct	39	56
CL50	31 Jul	31 Jul	15 Oct	6 Nov	23	23.7
CL5	30 Jul	6 Aug	16 Oct	17 Oct	53	67.4

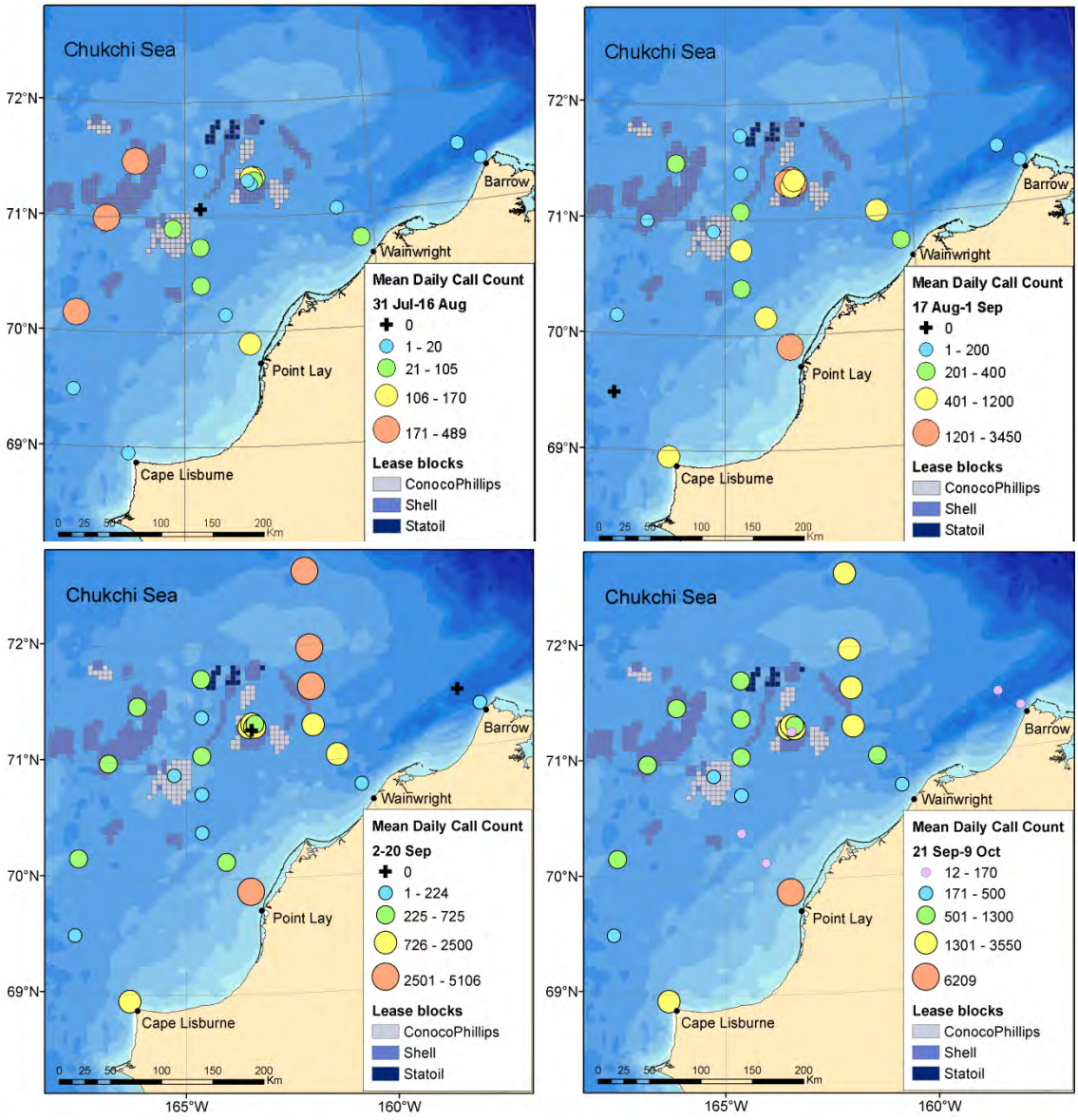

Figure C-13. Mean daily walrus call counts in the northeastern Chukchi Sea from (top left) 31 July to 16 Aug, (top right) 17 Aug to 1 Sep, (bottom left) 2 to 20 Sep, and (bottom right) 21 Sep to 9 Oct 2014.

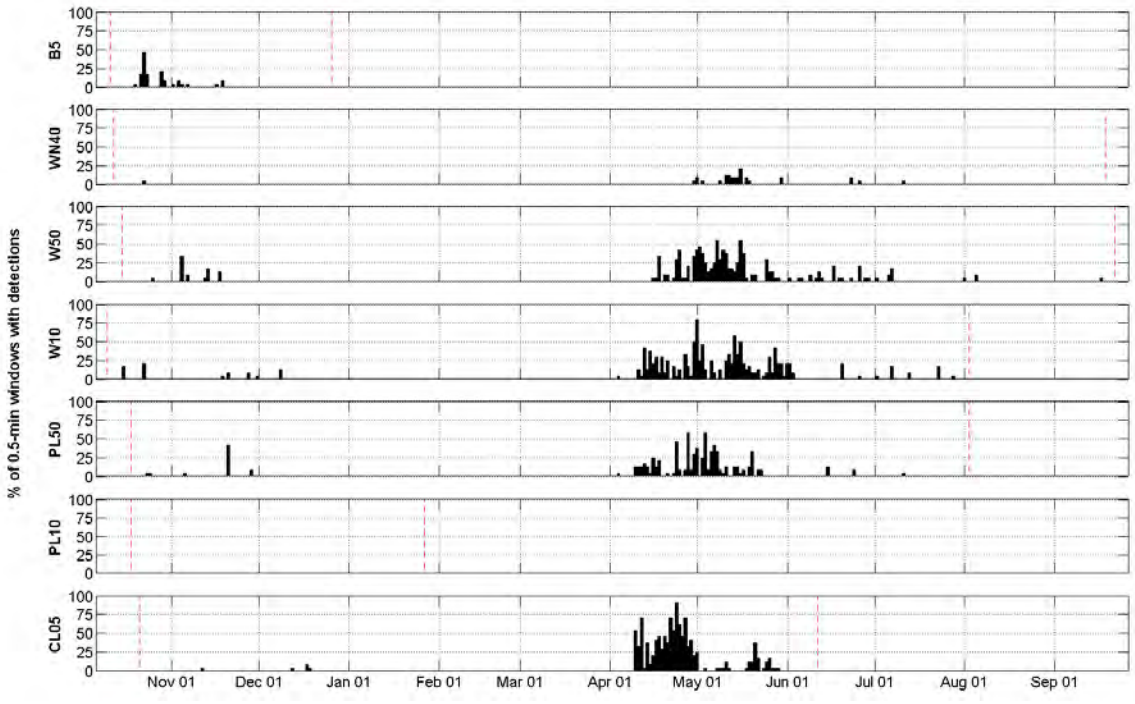

Figure C-14. Winter 2013–2014 daily beluga call detections: Daily proportion of sound files with call detections based on the manual analysis of 5% of the acoustic data recorded early September 2012 through late August 2013 in the northeastern Chukchi Sea for each station. Forty-eight sound files were recorded daily. Vertical dashed lines indicate recording start and end. Stations without call detections were omitted.

Table C-5. Winter 2013–2014 beluga whale call detections: Dates of first and last call detections, both possible (i.e., record start and end) and actual, and the number of days on which a call was detected manually for each recording station in the northeastern Chukchi Sea. The recorders operated for 5 min every 30 min.

Station	Record start	Fall 2013			Spring 2014			Record end
		First detection	Last detection	Detection days	First detection	Last detection	Detection days	
B05	11 Oct	19 Oct	18 Nov	12	--	--	0	26 Dec
WN40	12 Oct	22 Oct	22 Oct	1	30 Apr	11 Jul	15	18 Sep
W50	15 Oct	25 Oct	18 Nov	6	16 Apr	16 Sep	54	21 Sep
W10	11 Oct	15 Oct	8 Dec	9	4 Apr	28 Jul	56	2 Aug
PL50	19 Oct	24 Oct	28 Nov	5	4 Apr	11 Jul	36	2 Aug
CL05	20 Oct	12 Nov	18 Dec	3	9 Apr	28 May	36	11 Jun

-- No detections.

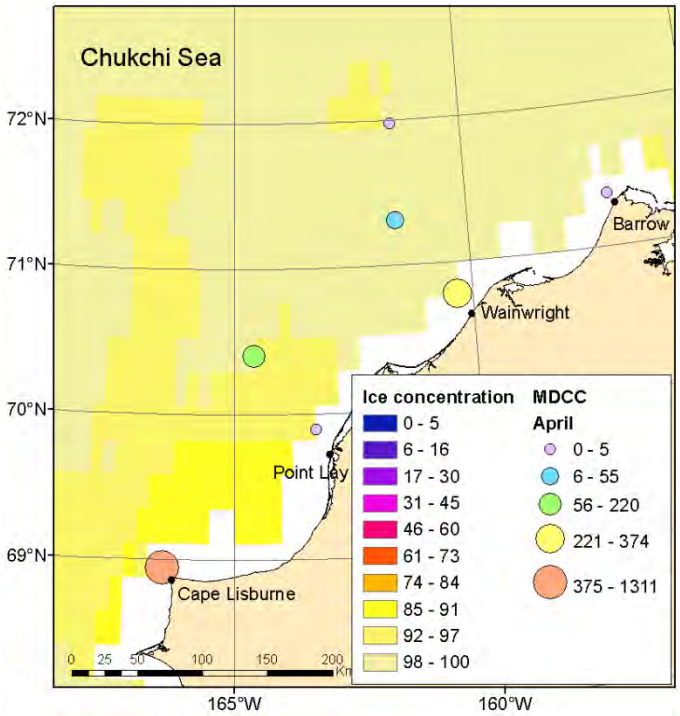

Figure C-15. Beluga whale call count estimates* in the Chukchi Sea for April 2013 at all winter 2012-2013 recorders. Areas of complete ice coverage are shown in gray for 15 Apr 2013 (NOAA 2008).

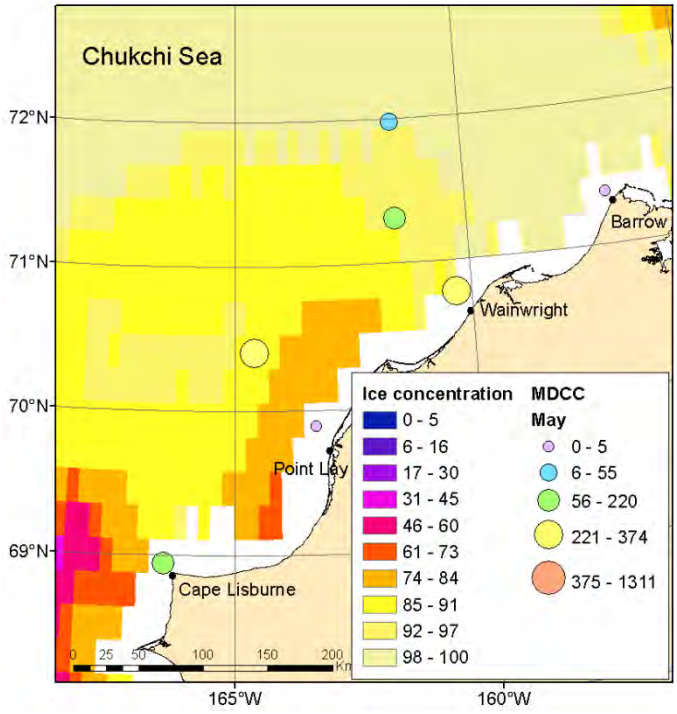

Figure C-16. Beluga whale call count estimates* in the Chukchi Sea for May 2013 at all winter 2012-2013 recorders. Areas of complete ice coverage are shown in gray for 15 May 2013 (NOAA 2008).

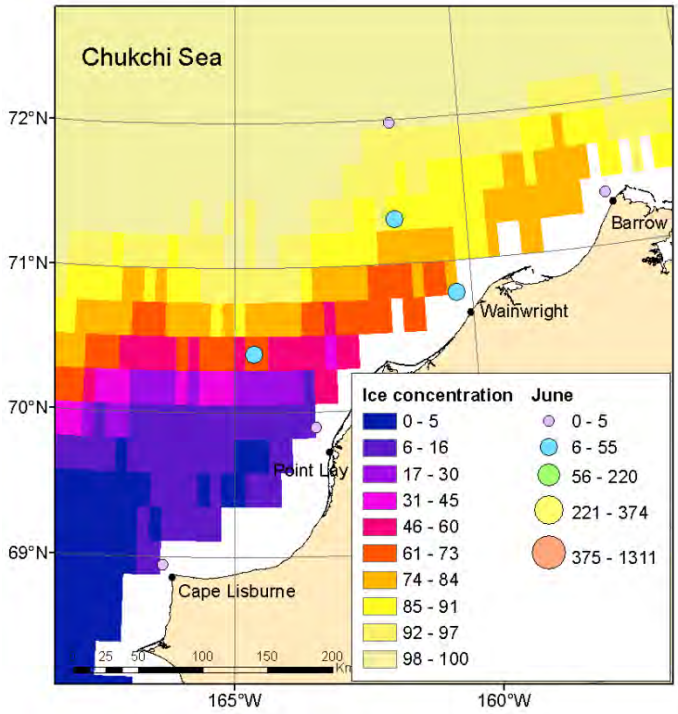

Figure C-17. Beluga whale call count estimates* in the Chukchi Sea for June 2013 at all winter 2012–2013 recorders. Areas of complete ice coverage are shown in gray for 15 Jun 2013 (NOAA 2008). The blue background indicates open water.

Table C-6. Summer 2014 beluga call detections: Dates of first and last call detections, both possible (i.e., record start and end) and actual, and the percent of days on which a call was detected for each recording station in the northeastern Chukchi Sea. Stations without call detections are omitted.

Station	Record start	First detection	Last detection	Record end	Detection days	% Days with detection
BGB	5 Aug	17 Aug	12 Oct	14 Oct	5	7.2
BGC	5 Aug	9 Aug	9 Oct	15 Oct	13	18.3
BGE	5 Aug	21 Aug	11 Oct	14 Oct	8	11.4
BGF	5 Aug	20 Aug	9 Oct	14 Oct	6	8.6
B15	6 Aug	7 Aug	13 Nov	13 Nov	34	34.7
B5	6 Aug	6 Aug	8 Nov	9 Nov	46	48.9
WN80	2 Sep	2 Oct	8 Oct	12 Oct	3	7.3
WN40	2 Sep	6 Sep	8 Oct	12 Oct	7	17
WN20	2 Sep	21 Sep	8 Oct	16 Oct	4	9.1
W50	2 Sep	16 Sep	13 Oct	16 Oct	12	27.3
W30	6 Aug	9 Aug	12 Oct	15 Oct	20	28.5
W10	3 Aug	27 Aug	8 Oct	15 Oct	8	10.9
PLN80	21 Aug	24 Aug	9 Oct	11 Oct	8	15.8
PLN60	5 Aug	9 Aug	9 Oct	13 Oct	4	5.8
PLN40	4 Aug	7 Oct	9 Oct	13 Oct	3	4.3
KL01	1 Aug	1 Oct	9 Oct	10 Oct	5	7.2
PLN20	1 Aug	1 Oct	8 Oct	10 Oct	6	8.6
PL50	2 Aug	7 Oct	7 Oct	9 Oct	1	1.5
CLN120B	31 Jul	10 Aug	10 Oct	11 Oct	14	19.7
CLN90B	31 Jul	18 Sep	10 Oct	10 Oct	6	8.4
CLN40	31 Jul	1 Oct	6 Oct	8 Oct	5	7.2
CL50	31 Jul	8 Oct	31 Oct	6 Nov	3	3.1

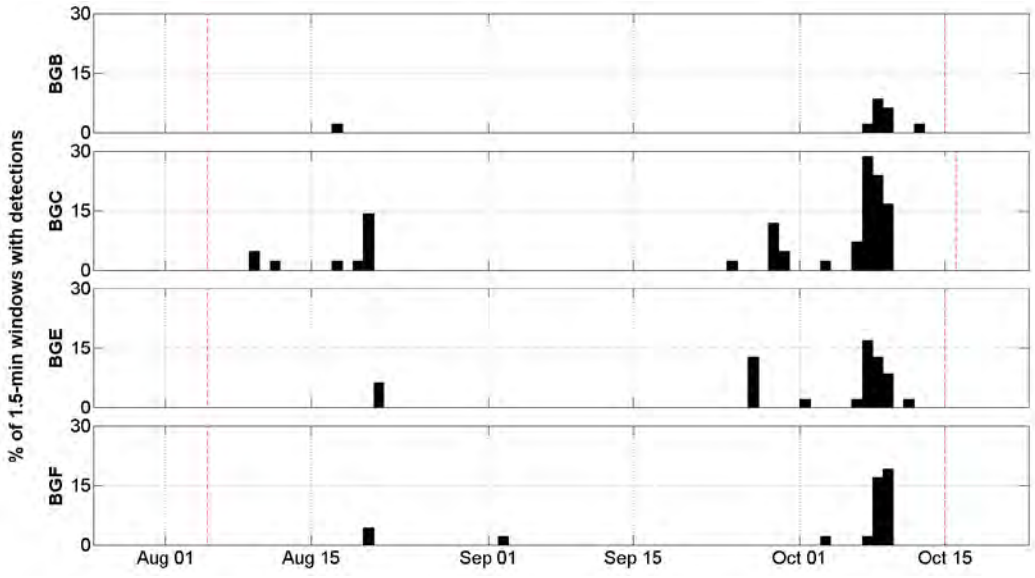

Figure C-18. Summer 2014 daily beluga call detections in the Burger lease area: Daily proportion of sound files with detections based on the manual analysis of 5% of the acoustic data recorded late July through mid-October 2014. Forty-eight sound files were recorded each day. Vertical dashed lines indicate recording start and end. Stations are ordered northeast (top) to southwest (bottom). Stations without call detections were omitted.

Table C-7. Winter 2013–2014 bearded seal call detections: Dates of first and last call detections, both possible (i.e., record start and end) and actual, and the number of days on which a call was detected manually for each recording station in the northeastern Chukchi Sea. The recorders operated for 5 min every 30 min.

Station	Record start	First detection	Last detection	Record end	Detection days
B05	11 Oct	19 Oct	26 Dec	26 Dec	22
WN40	12 Oct	16 Oct	30 Jun	18 Sep	194
W50	15 Oct	22 Oct	8 Sep	21 Sep	226
W10	11 Oct	12 Oct	29 Jun	2 Aug	201
PL50	19 Oct	22 Oct	20 Jun	2 Aug	183
PL10	18 Oct	22 Oct	27 Jan	27 Jan	43
CL05	20 Oct	23 Oct	11 Jun	11 Jun	150

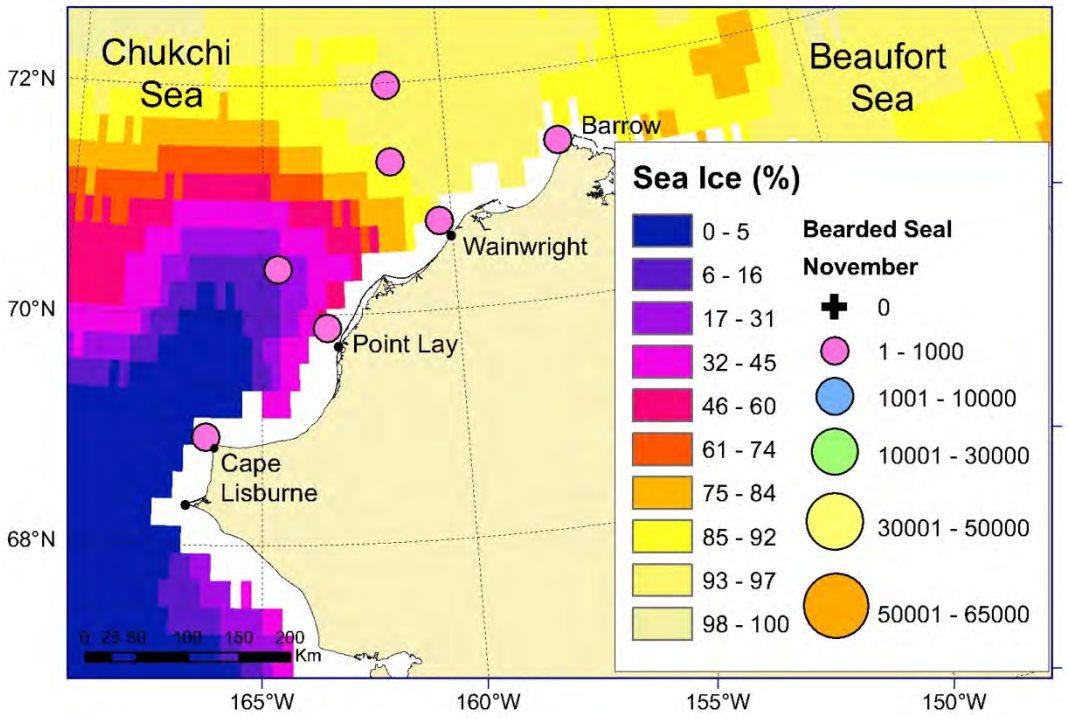

Figure C-19. Bearded seal call count estimates* in the Chukchi Sea for November 2013 at all winter 2013–2014 recorders.

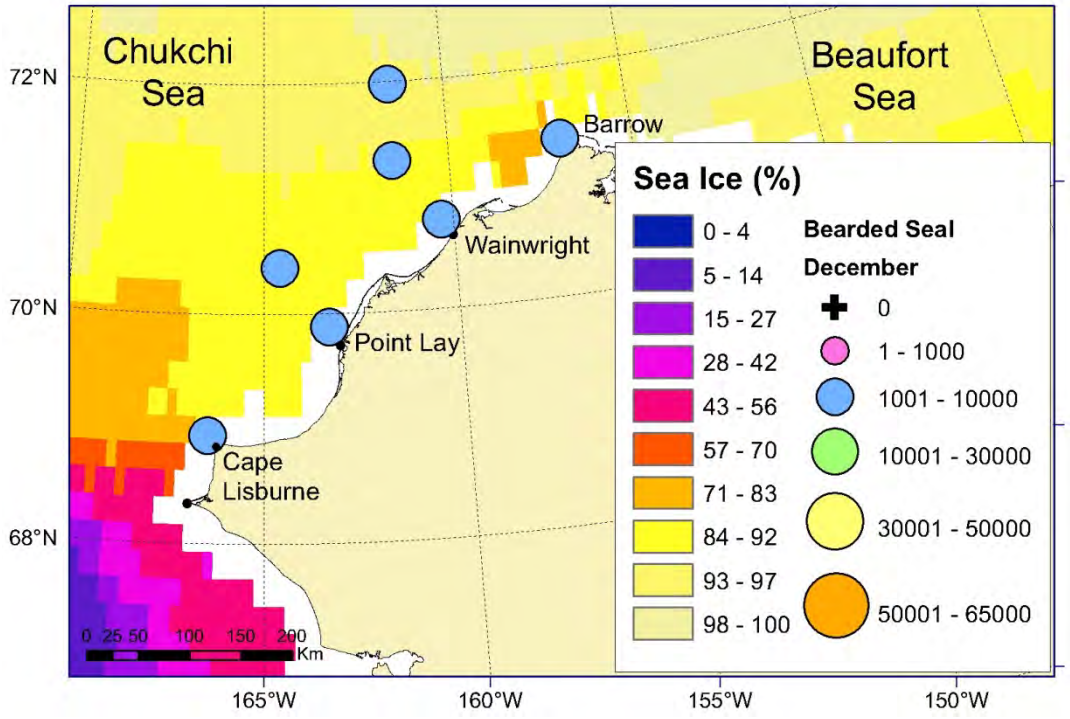

Figure C-20. Bearded seal call count estimates* in the Chukchi Sea for December 2013 at all winter 2013–2014 recorders.

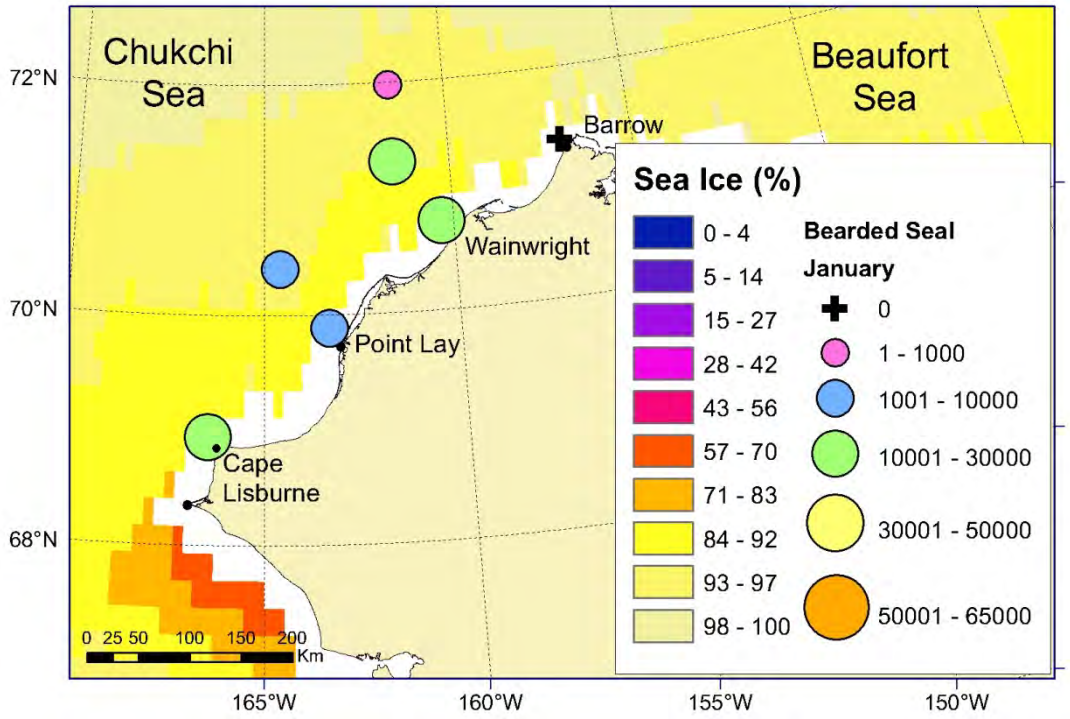

Figure C-21. Bearded seal call count estimates* in the Chukchi Sea for January 2014 at all winter 2013–2014 recorders.

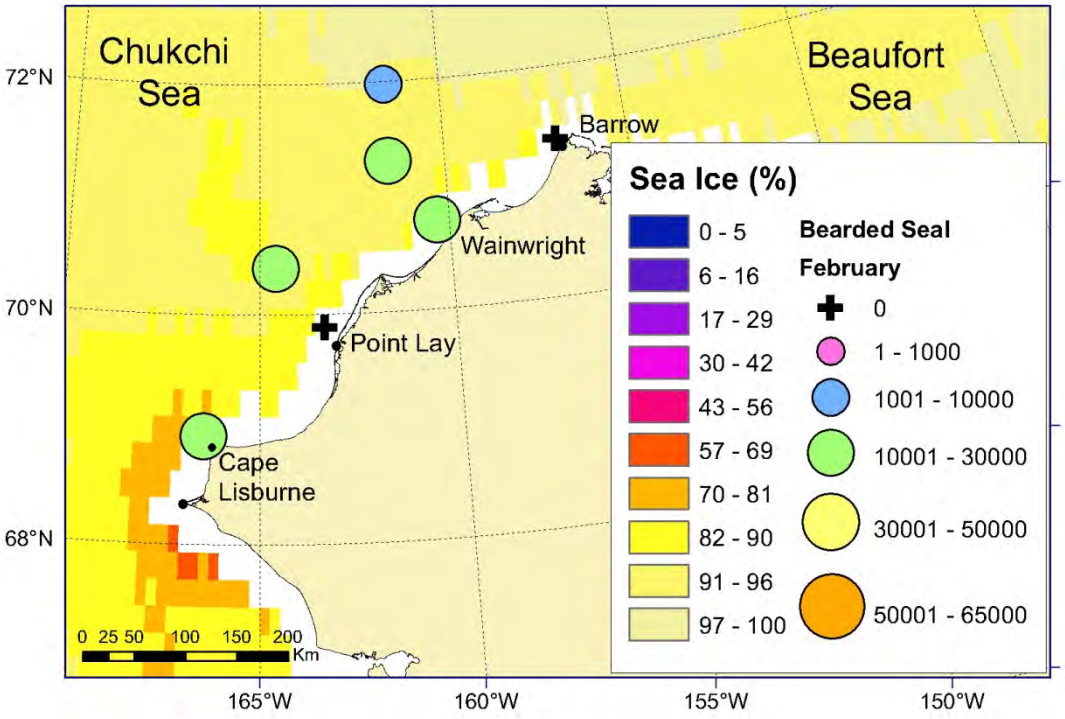

Figure C-22. Bearded seal call count estimates* in the Chukchi Sea for February 2014 at all winter 2013–2014 recorders.

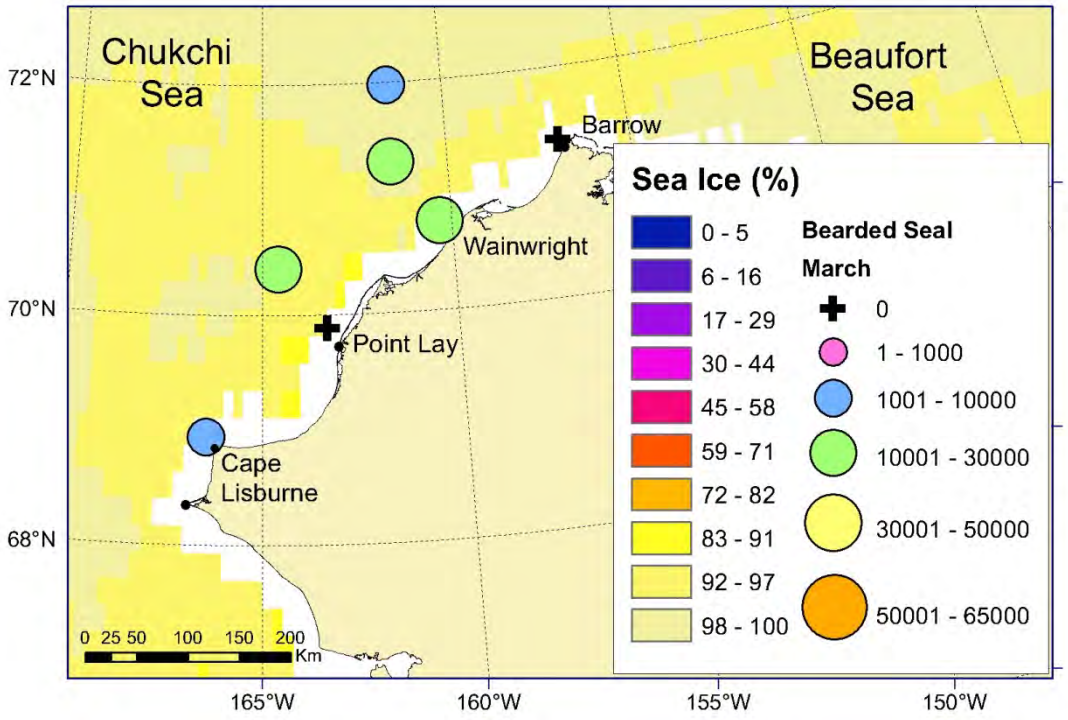

Figure C-23. Bearded seal call count estimates* in the Chukchi Sea for March 2014 at all winter 2013–2014 recorders.

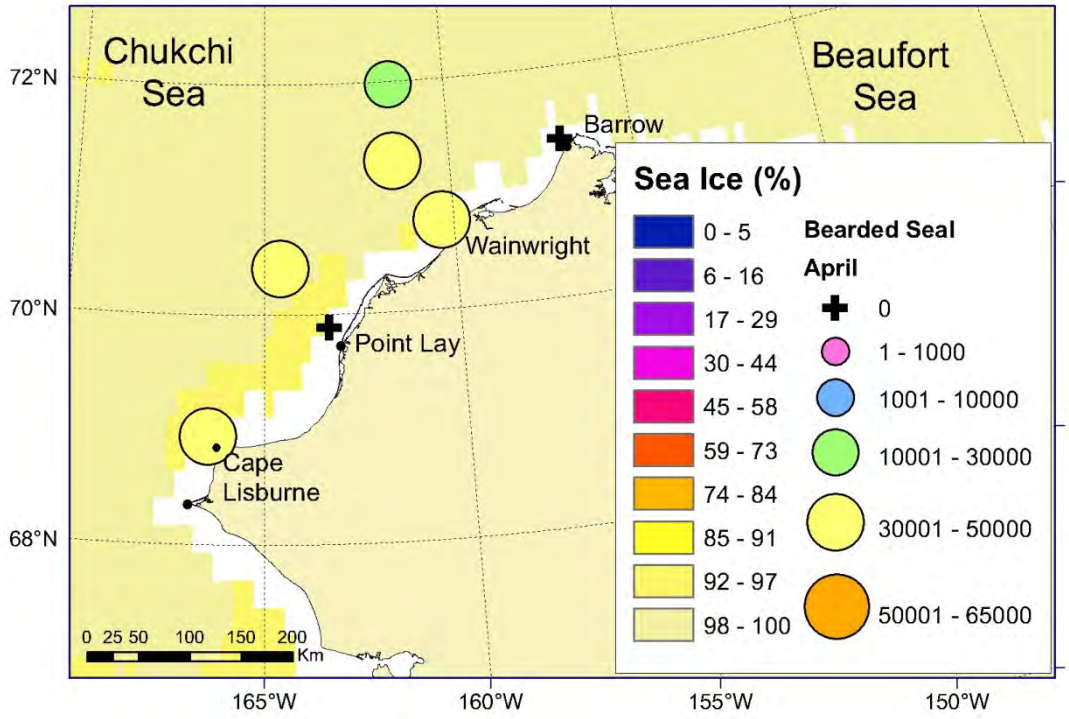

Figure C-24. Bearded seal call count estimates* in the Chukchi Sea for April 2014 at all winter 2013–2014 recorders.

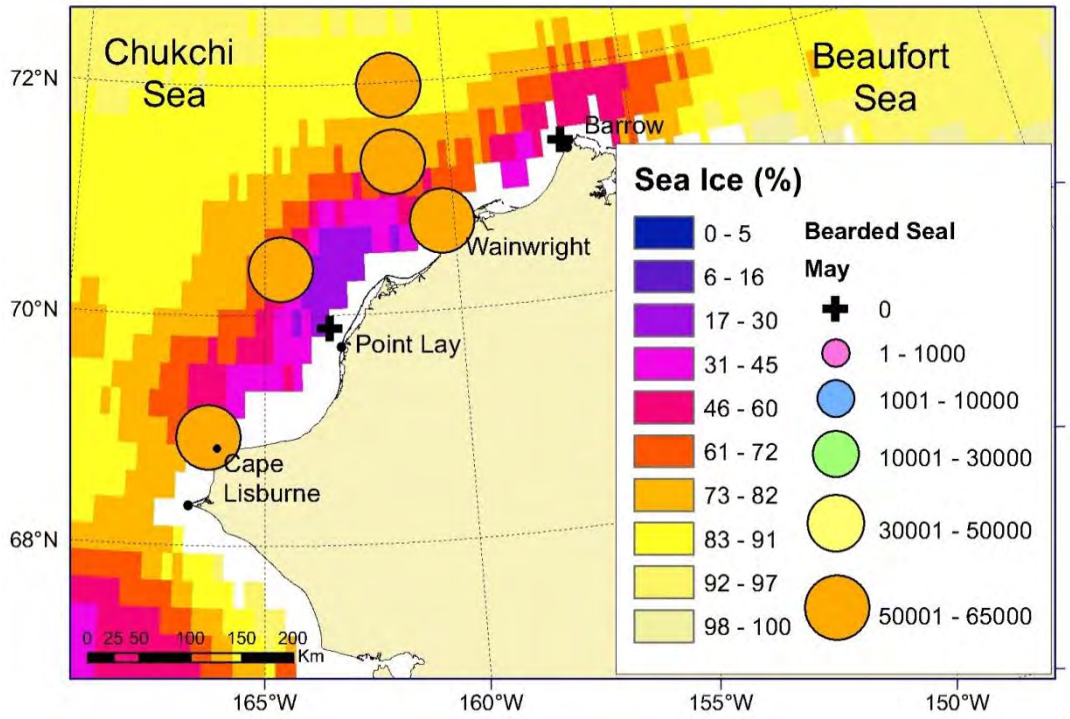

Figure C-25. Bearded seal call count estimates* in the Chukchi Sea for May 2014 at all winter 2013–2014 recorders.

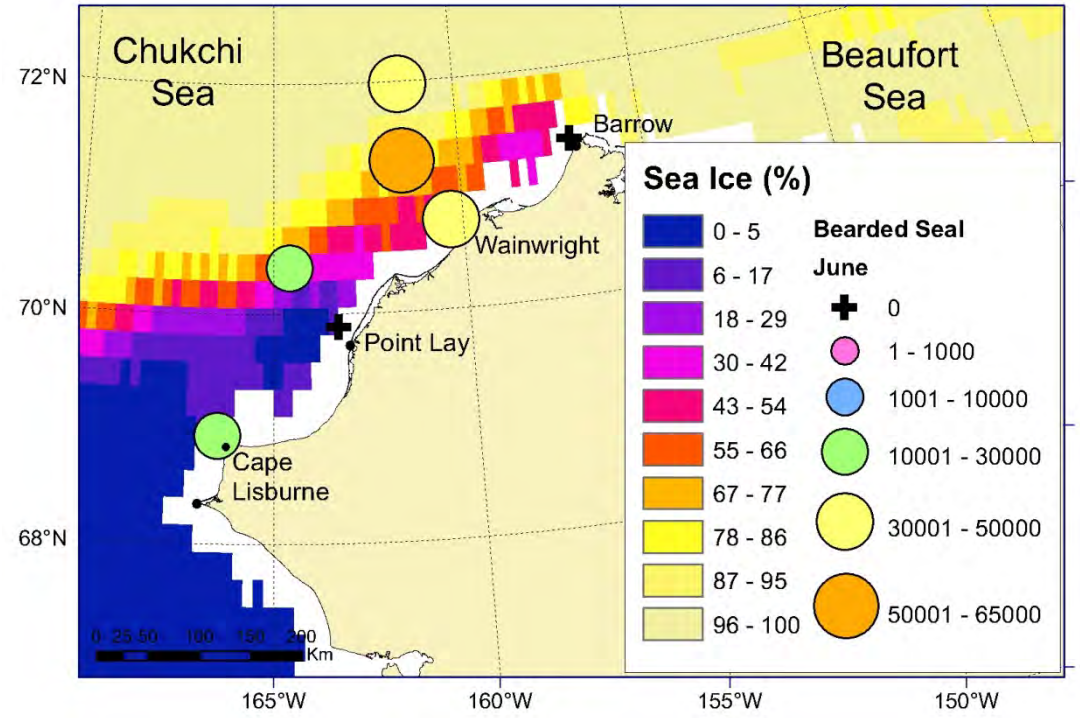

Figure C-26. Bearded seal call count estimates* in the Chukchi Sea for June 2014 at all winter 2013–2014 recorders.

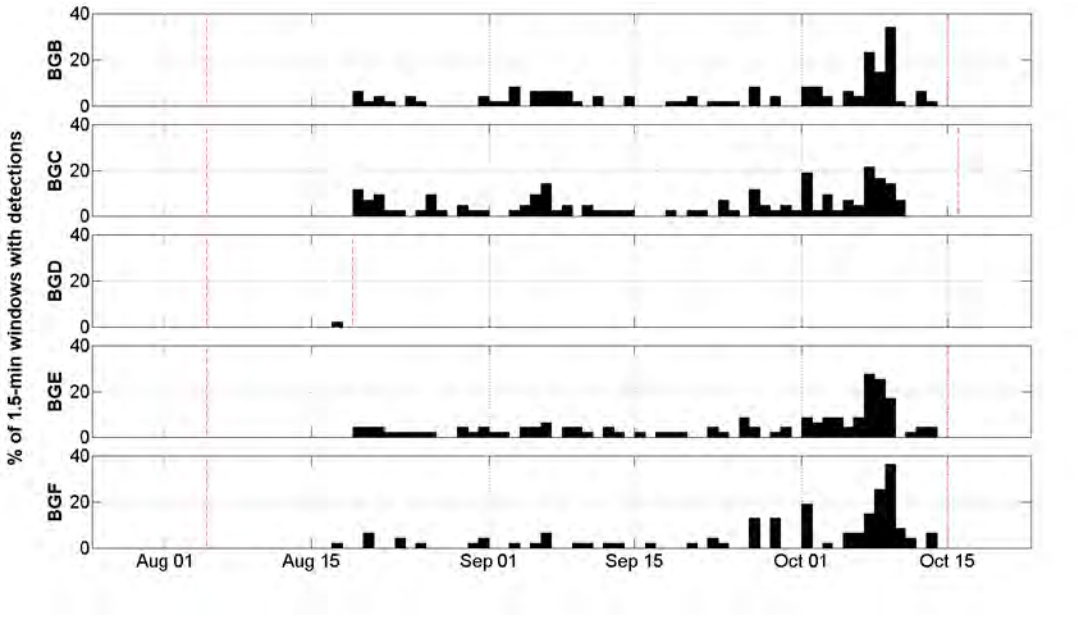

Figure C-27. Summer 2014 daily bearded seal call detections: Daily proportion of sound files with detections based on the manual analysis of 5% of the acoustic data recorded late July through mid-October 2014 in the Burger study area. Forty-eight sound files were recorded daily. Vertical dashed lines indicate recording start and end. **Error! Reference source not found.**Stations without call detections were omitted.

Table C-8. Summer 2014 bearded seal call detections: Dates of first and last call detections, both possible (i.e., record start and end) and actual, and the percent of days on which a call was detected for each recording station in the northeastern Chukchi Sea. Stations without call detections are omitted.

Station	Record start	First detection	Last detection	Record end	Detection days	% Days with detection
BGB	5 Aug	19 Aug	13 Oct	14 Oct	36	51.6
BGC	5 Aug	19 Aug	10 Oct	15 Oct	42	59.2
BGD	5 Aug	17 Aug	17 Aug	18 Aug	1	7.4
BGE	5 Aug	19 Aug	13 Oct	14 Oct	43	61.4
BGF	5 Aug	17 Aug	13 Oct	14 Oct	29	41.5
B15	6 Aug	10 Aug	8 Nov	13 Nov	52	54.1
B5	6 Aug	15 Aug	8 Nov	9 Nov	41	43.6
WN80	2 Sep	3 Sep	9 Oct	12 Oct	11	26.8
WN40	2 Sep	8 Sep	12 Oct	12 Oct	25	61
WN20	2 Sep	3 Sep	12 Oct	16 Oct	29	66
W50	2 Sep	2 Sep	15 Oct	16 Oct	32	72.9
W30	6 Aug	8 Aug	15 Oct	15 Oct	32	45.6
W10	3 Aug	3 Aug	15 Oct	15 Oct	38	51.6
PLN80	21 Aug	26 Aug	10 Oct	11 Oct	20	39.4
PLN60	5 Aug	19 Aug	12 Oct	13 Oct	31	45
PLN40	4 Aug	17 Aug	12 Oct	13 Oct	25	36
KL01	1 Aug	17 Aug	8 Oct	10 Oct	6	8.6
PLN20	1 Aug	9 Aug	9 Oct	10 Oct	38	54.8
PL50	2 Aug	14 Aug	7Oct	9 Oct	11	16.0
PL30	2 Aug	18 Aug	6 Sep	9 Oct	4	5.8
PL10	2 Aug	26 Sep	8 Oct	9 Oct	2	2.9
CLN120B	31 Jul	26 Aug	9 Oct	11 Oct	16	22.5
CLN90B	31 Jul	15 Aug	9 Oct	10 Oct	9	12.6
CLN40	31 Jul	1 Aug	6 Oct	8 Oct	12	17.2
CL50	31 Jul	26 Sep	3 Nov	6 Nov	16	16.5

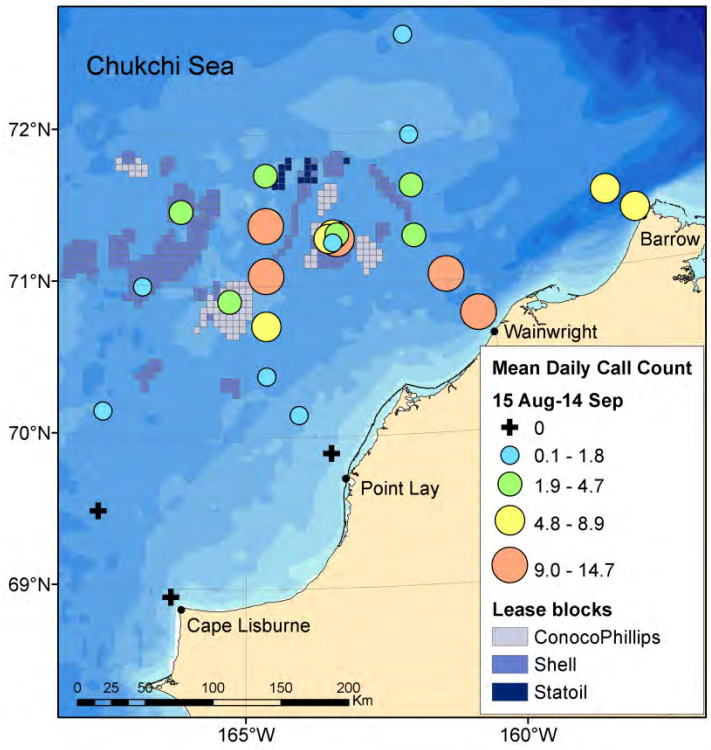

Figure C-28. Mean daily bearded seal call counts in the northeastern Chukchi Sea from 15 Aug to 14 Sep 2014 at all operational summer recording stations in the northeastern Chukchi Sea.

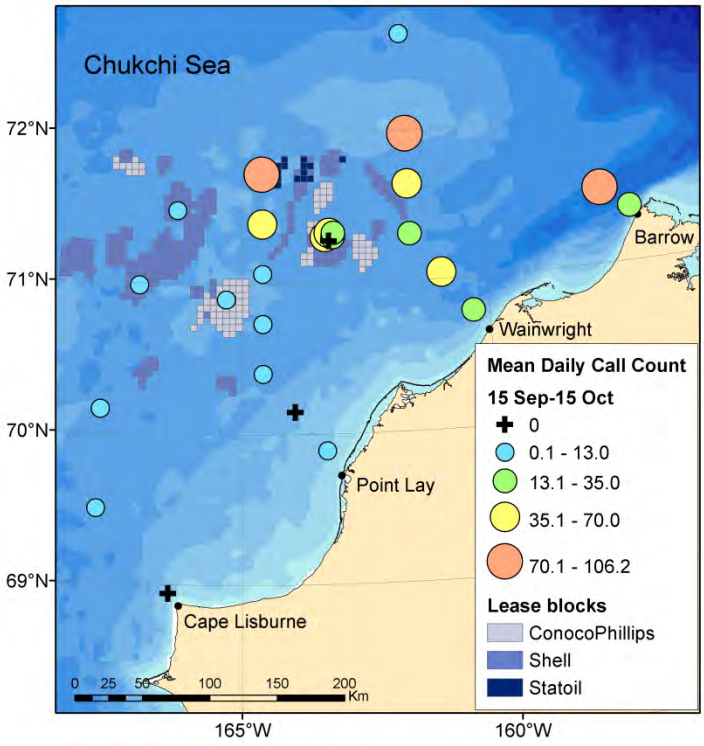

Figure C-29. Mean daily bearded seal call counts in the northeastern Chukchi Sea from 15 Sep to 15 Oct 2014 at all operational summer recording stations in the northeastern Chukchi Sea.

Table C-9. Summer 2014 gray whale call detections: Dates of first and last call detections, both possible (i.e., record start and end) and actual, and the percent of days on which a call was detected for each recording station in the northeastern Chukchi Sea. Stations without call detections are omitted.

Station	Record start	First detection	Last detection	Record end	Detection days	% Days with detection
B5	6 Aug	7 Aug	4 Sep	9 Nov	10	10.6
WN40	2 Sep	6 Sep	6 Sep	12 Oct	1	2.4
W50	2 Sep	18 Sep	18 Sep	16 Oct	1	2.3
WN20	2 Sep	6 Sep	1 Oct	16 Oct	4	9.1
W30	6 Aug	11 Aug	14 Oct	15 Oct	7	10
W10	3 Aug	3 Aug	2 Oct	15 Oct	26	35.3
BGC	5 Aug	28 Aug	16 Sep	15 Oct	8	11.1
BGD	5 Aug	6 Aug	6 Aug	18 Aug	1	7.7
BGE	5 Aug	6 Sep	6 Sep	14 Oct	1	1.5
BGF	5 Aug	18 Aug	18 Aug	14 Oct	1	1.5
PLN80	21 Aug	25 Aug	8 Sep	11 Oct	3	5.9
PLN60	5 Aug	18 Aug	10 Oct	13 Oct	11	16
PLN40	4 Aug	29 Aug	7 Sep	13 Oct	2	2.9
KL01	1 Aug	17 Aug	23 Aug	10 Oct	3	4.3
PLN20	1 Aug	18 Aug	22 Sep	10 Oct	4	5.8
PL50	2 Aug	6 Aug	7 Oct	9 Oct	8	11.7
PL30	2 Aug	2 Aug	5 Sep	9 Oct	2	2.9
CLN120B	31 Jul	3 Aug	8 Sep	11 Oct	2	2.8
CLN90B	31 Jul	5 Sep	5 Oct	10 Oct	3	4.2
CLN40	31 Jul	6 Aug	20 Sep	8 Oct	3	4.3
CL50	31 Jul	31 Jul	4 Sep	6 Nov	2	2.1
CL5	30 Jul	26 Sep	26 Sep	17 Oct	1	1.3

Table C-10. Summer 2014 killer whale call detections: Dates of first and last call detections, both possible (i.e., record start and end) and actual, and the percent of days on which a call was detected for each recording station in the northeastern Chukchi Sea. Stations without call detections are omitted.

Station	Record start	First detection	Last detection	Record end	Detection days	% Days with detection
BGB	5 Aug	15 Sep	7 Oct	14 Oct	2	2.9
BGC	5 Aug	26 Aug	15 Sep	15 Oct	3	4.2
BGE	5 Aug	8 Oct	8 Oct	14 Oct	1	1.4
W30	6 Aug	19 Aug	27 Aug	15 Oct	2	2.8
PLN60	5 Aug	19 Aug	20 Aug	13 Oct	2	2.9
PLN40	4 Aug	20 Aug	20 Aug	13 Oct	1	1.4
KL01	1 Aug	20 Aug	7 Oct	10 Oct	2	2.9
PLN20	1 Aug	20 Aug	20 Aug	10 Oct	1	1.4
PL50	2 Aug	21 Aug	7 Oct	9 Oct	2	2.9
PL10	2 Aug	8 Oct	8 Oct	9 Oct	1	1.5
CLN120B	31 Jul	19 Aug	19 Aug	11 Oct	1	1.4
CLN90B	31 Jul	19 Aug	7 Oct	10 Oct	3	4.2
CL50	31 Jul	17 Aug	8 Sep	6 Nov	2	2.1
CL5	30 Jul	19 Aug	19 Aug	17 Oct	1	1.3

Table C-11. Winter 2013–2014 ringed seal call detections: Dates of first and last call detections, both possible (i.e., record start and end) and actual, and the percent of days on which a call was detected for each recording station in the northeastern Chukchi Sea. Stations without call detections are omitted. The recorders operated for 30–40 min every 4 h.

Station	Record start	First detection	Last detection	Record end	Detection days
B5	11 Oct	23 Nov	23 Nov	26 Dec	1
WN40	12 Oct	23 Dec	13 Jun	18 Sep	39
W50	15 Oct	30 Oct	7 Jun	21 Sep	127
W10	11 Oct	31 Oct	9 May	2 Aug	16
PL50	19 Oct	31 Oct	13 Jun	2 Aug	125
PL10	19 Oct	4 Nov	23 Dec	27 Jan	2
CL05	21 Oct	6 Nov	26 May	11 Jun	95

Table C-12. Summer 2014 ringed seal call detections: Dates of first and last call detections, both possible (i.e., record start and end) and actual, and the percent of days on which a call was detected for each recording station in the northeastern Chukchi Sea. Stations without call detections are omitted.

Station	Record start	First detection	Last detection	Record end	Detection days	% Days with detection
BGC	05 Aug	19 Aug	26 Aug	15 Oct	4	5.6
BGF	05 Aug	07 Aug	07 Aug	14 Oct	1	1.4
B15	06 Aug	27 Aug	13 Nov	13 Nov	13	13.3
B5	06 Aug	18 Aug	28 Oct	09 Nov	4	4.3
WN80	02 Sep	08 Oct	08 Oct	12 Oct	1	2.5
W30	06 Aug	21 Aug	26 Sep	15 Oct	3	4.3
W10	03 Aug	18 Aug	22 Sep	15 Oct	4	5.5
PLN60	05 Aug	06 Oct	09 Oct	13 Oct	2	2.9
PLN40	04 Aug	18 Aug	07 Oct	13 Oct	2	2.9
KL01	01 Aug	03 Aug	06 Sep	10 Oct	4	5.7
PLN20	01 Aug	03 Aug	22 Sep	10 Oct	8	11.4
PL50	02 Aug	03 Aug	30 Sep	09 Oct	4	5.9
PL30	02 Aug	18 Aug	21 Aug	09 Oct	4	5.9
CLN120B	31 Jul	19 Aug	26 Aug	11 Oct	3	4.2
CLN90B	31 Jul	21 Aug	21 Aug	10 Oct	1	1.4
CLN40	31 Jul	05 Sep	28 Sep	08 Oct	2	2.9
CL50	31 Jul	06 Aug	30 Sep	06 Nov	4	4.1

Appendix D. Estimating the Detection Range of Bowhead Moans

This appendix describes how the detection range of bowhead moans was calculated for each recorder of the summer 2014 Program.

D.1. Methods

The received sound level (*RL*) of a bowhead moan at a recorder is defined by the following equation (Urick 1983):

$$RL = SL–TL \quad (1)$$

where *SL* is the source level of the bowhead moan, and *TL* is the transmission loss between the whale and the hydrophone. The detection range of a bowhead moan was assumed to be the distance from the recorder for which the received level of the bowhead call equaled or exceeded the noise level at the recorder (*NL*):

$$NL = RL. \quad (2)$$

Cummings and Holliday (1987) and Clark et al. (1986) estimated that source levels of simple moans range from ~128 to 178 dB re 1 µPa at 1 m. MacDonnell et al. (2011) estimated that bowhead moans recorded near the Burger lease area had source levels of 144.3 ± 4.6 dB re 1 µPa at 1 m (mean ± standard deviation), with minimum and maximum levels of 129.7 and 164.4 dB respectively (Figure D-1). These latter values were used for estimating the bowhead detection range at each recorder Equation 1.

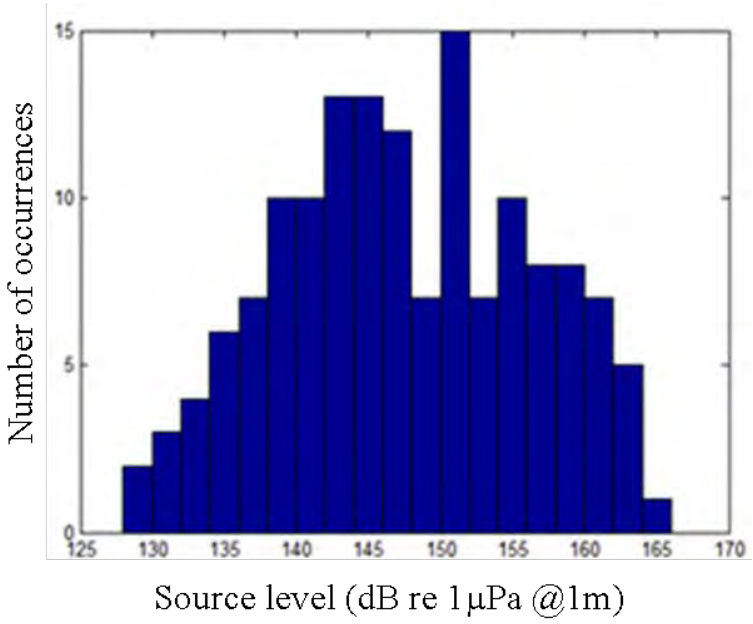

Figure D-1. Distribution of source levels reported by MacDonnell and Martin (2011).

Transmission loss values used for estimating the bowhead detection range came from a previous study by MacDonnell and Martin (2011) at the Burger lease area. In that study, transmission loss was calculated between 89 and 447 Hz using JASCO's Marine Operations Noise Model (Hannay and Racca 2005, Austin 2012). This frequency range comes from using the seven 1/3-octave-bands centered between 100 and 400 Hz. Figure D-2 shows a transmission loss map calculated by (MacDonnell and Martin 2011) at BG01 (summer 2009).

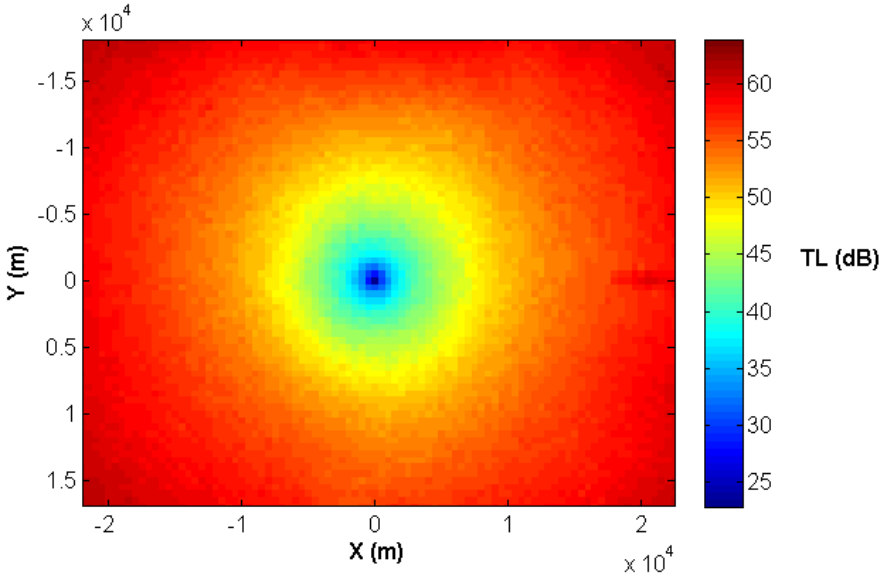

Figure D-2. Map of the transmission loss values calculated by MacDonnell and Martin (2011) at station BG01 (summer 2009). x is east-west distance. y is north-south distance.

The water depth in the eastern Chukchi Sea is nearly constant. Consequently, the transmission loss is nearly the same for all azimuths (Figure D-2). To simplify the calculation of the detection range, the transmission loss values from MacDonnell and Martin (2011) were represented by one equation:

$$TL(R) = A \log_{10}(R) + \alpha R \quad (3)$$

where, R is the distance from hydrophone to whale, A is the spreading coefficient and α is the attenuation coefficient (Urick 1983). The coefficients A and α were defined by fitting (in the least square sense) Equation (3) to the average transmission loss taken in four different azimuths from the recorder (i.e., 0°, 90°, 180°, and 270°). Figure D-3(left) shows the transmission loss curve from MacDonnell et al. (2011) at the location BG01 in four different azimuths.
Figure D-3 (right) shows the average transmission loss curve with its simplified transmission loss function.

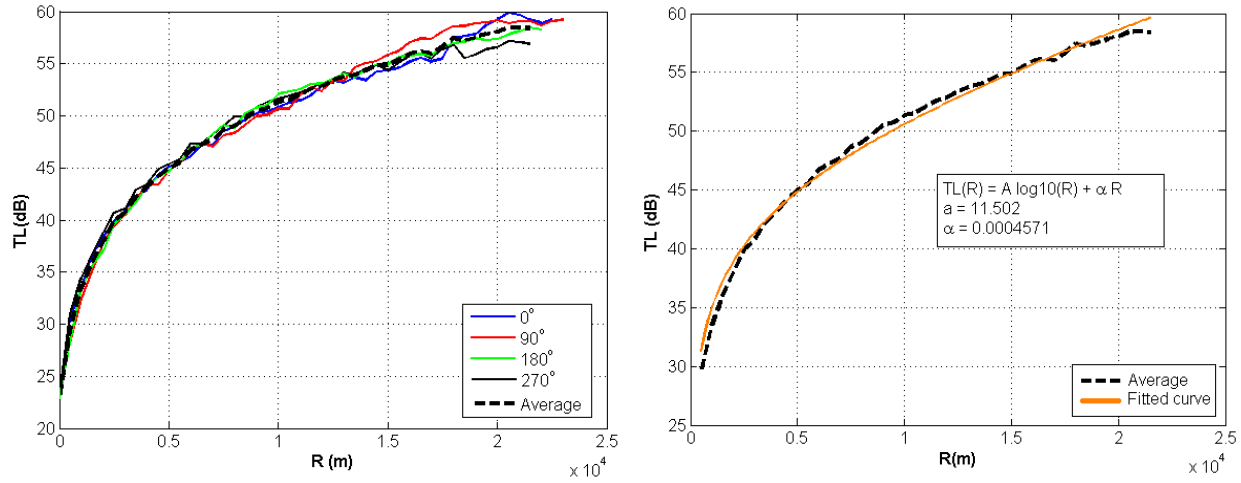

Figure D-3. Transmission loss modeled by MacDonnell et al. (2011) at location BG01. (Left) Transmission loss curves in four different azimuths. (Right) Average transmission loss and its simplified transmission loss function.

Coefficients A and α were calculated for each location modeled by MacDonnell et al. (2011) and then averaged to obtain a single set of coefficients for the whole area. Final coefficients used for the detection range analysis were $A = 11.29$ and $\alpha = 0.00057$.

Noise levels used for estimating the bowhead detection range were calculated for every minute of recording by summing the 1/3-octave-band levels between 89 and 447 Hz.

The detection range was calculated at each recorder and for each minute of recording. The probability of detecting a bowhead moan at a given range was the number of 1 min recordings with a detection range equal to or greater than the given range divided by the number of 1 min recordings. Detection ranges were calculated independently for each recorder.

A Monte Carlo method accounted for the measured variability in source levels. Detection ranges were re-calculated 50 times by randomly choosing 50 normally distributed source level values, with the means and standard deviations defined by MacDonnell et al. (2011; Figure D-3), Consequently, a distribution of probability is associated with each range.

D.2. Results

Figure D-4 through Figure D-30 show the detection probability of bowhead moans at each recorder.

The largest 80th percentile detection ranges occurred at Stations WN80, B15, and PLN20; 7.2 km, 6.1 km, and 5.9 km respectively. The smallest 80th percentile detection ranges occurred at Stations CL50, PL10, and PL30: 2.5 km, 2.6 km, and 2.7 km, respectively.

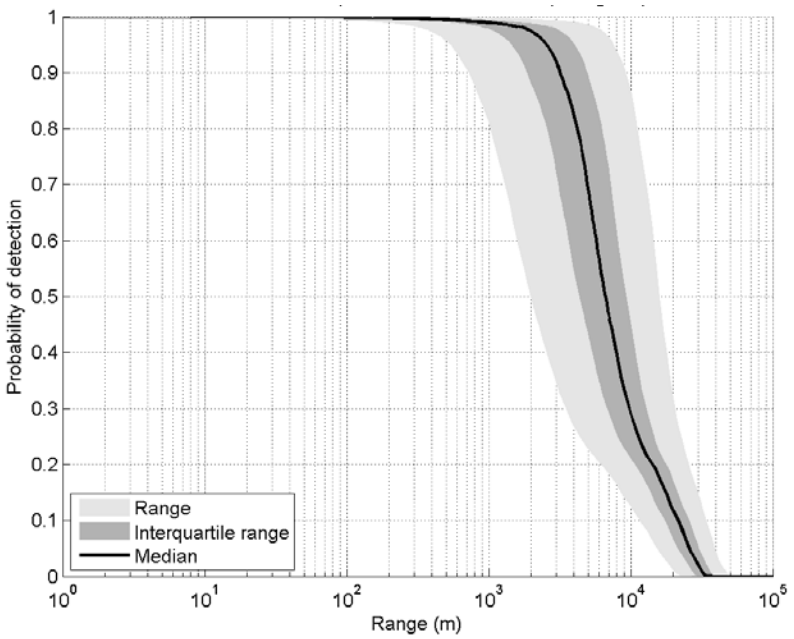

Figure D-4. Detection probability of bowhead moans at Station B05. The solid black line represents the median probability of detection. The light gray areas represent the probability range (from the 5th to the 95th percentile), and the dark gray areas the probability interquartile range (from the 25th to the 75th percentile).

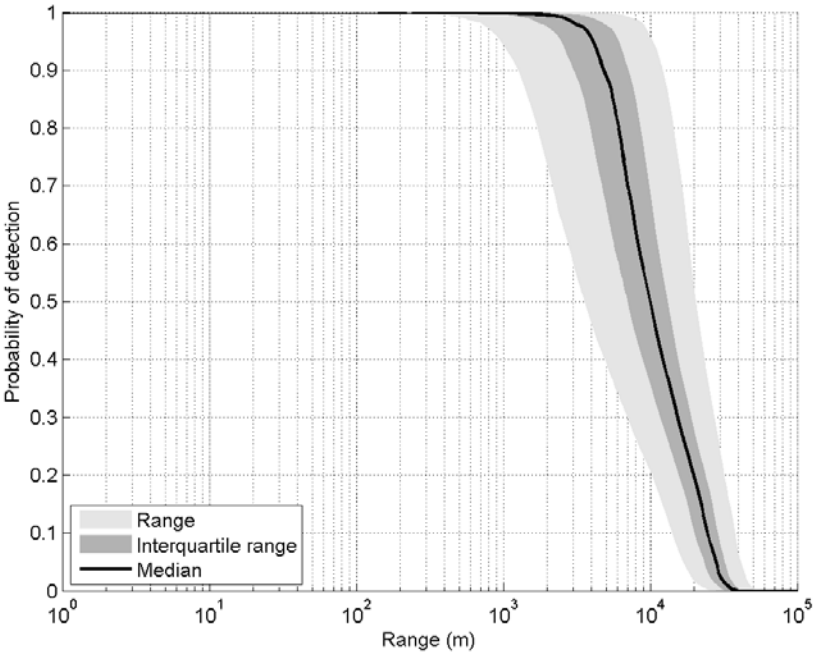

Figure D-5. Detection probability of bowhead moans at Station B15. The solid black line represents the median probability of detection. The light gray areas represent the probability range (from the 5th to the 95th percentile), and the dark gray areas the probability interquartile range (from the 25th to the 75th percentile).

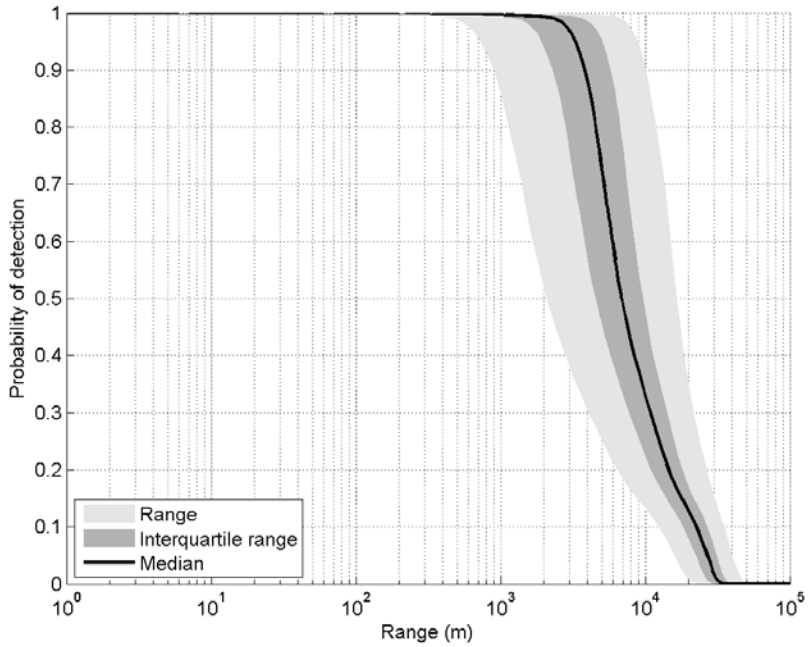

Figure D-6. Detection probability of bowhead moans at Station W10. The solid black line represents the median probability of detection. The light gray areas represent the probability range (from the 5th to the 95th percentile), and the dark gray areas the probability interquartile range (from the 25th to the 75th percentile).

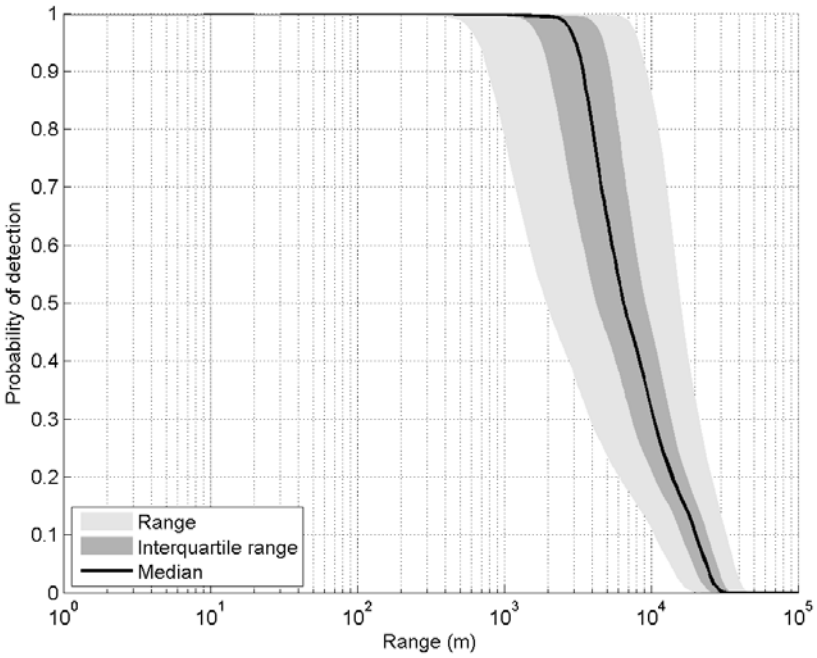

Figure D-7. Detection probability of bowhead moans at Station W30. The solid black line represents the median probability of detection. The light gray areas represent the probability range (from the 5th to the 95th percentile), and the dark gray areas the probability interquartile range (from the 25th to the 75th percentile).

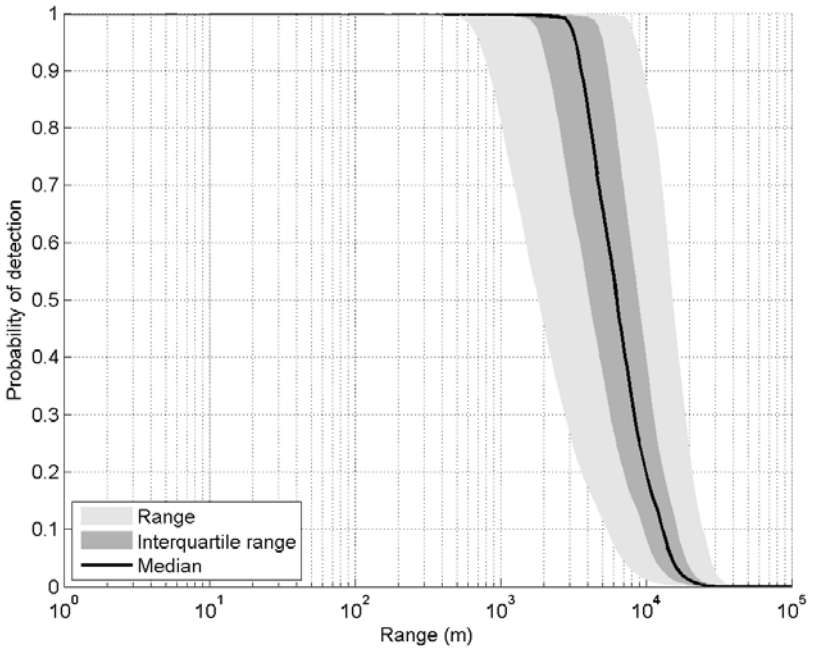

Figure D-8. Detection probability of bowhead moans at Station W50. The solid black line represents the median probability of detection. The light gray areas represent the probability range (from the 5th to the 95th percentile), and the dark gray areas the probability interquartile range (from the 25th to the 75th percentile).

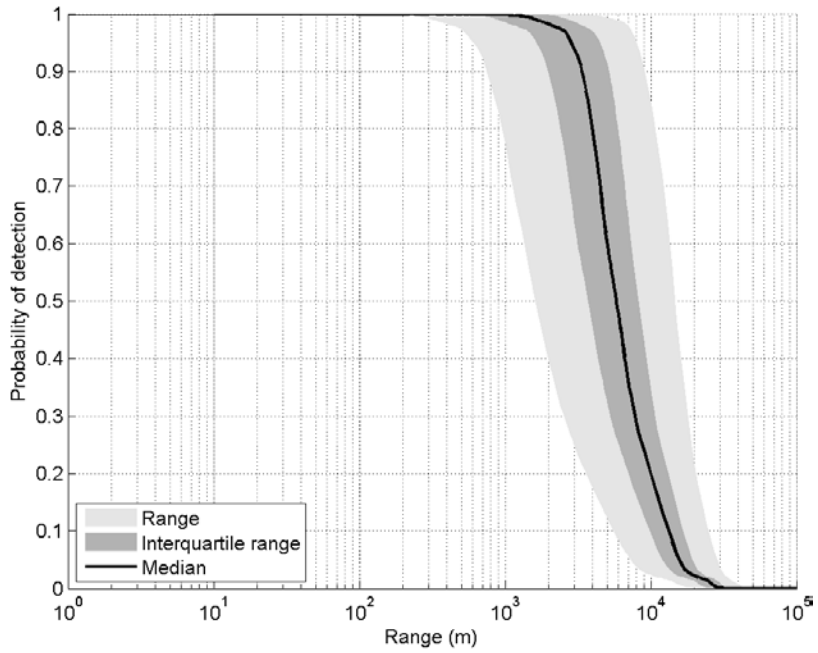

Figure D-9. Detection probability of bowhead moans at Station WN20. The solid black line represents the median probability of detection. The light gray areas represent the probability range (from the 5th to the 95th percentile), and the dark gray areas the probability interquartile range (from the 25th to the 75th percentile).

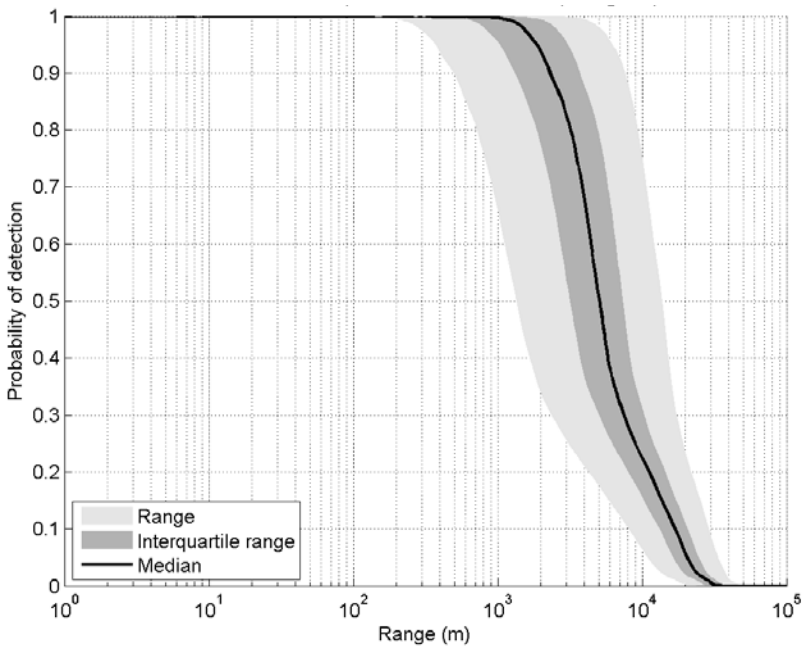

Figure D-10. Detection probability of bowhead moans at Station WN40. The solid black line represents the median probability of detection. The light gray areas represent the probability range (from the 5th to the 95th percentile), and the dark gray areas the probability interquartile range (from the 25th to the 75th percentile).

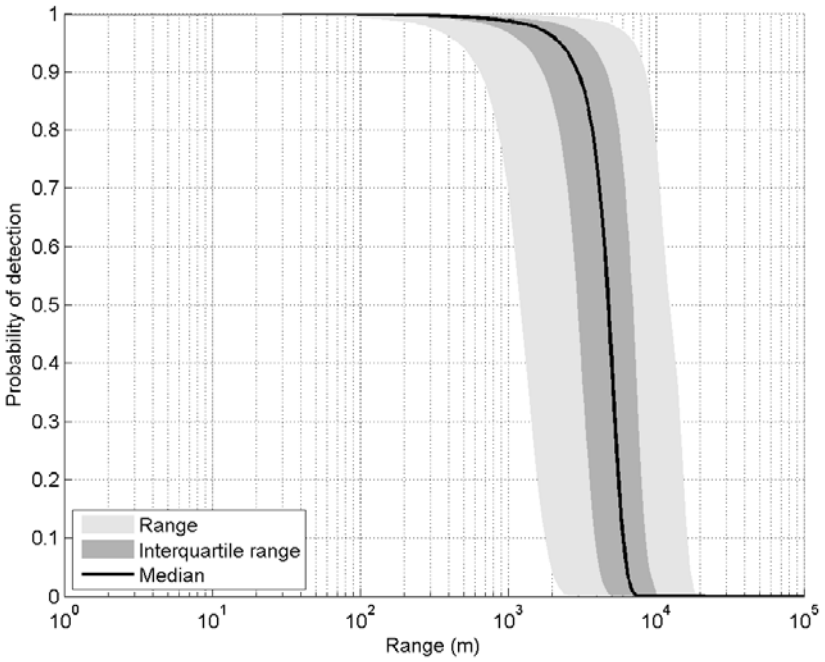

Figure D-11. Detection probability of bowhead moans at Station WN60. The solid black line represents the median probability of detection. The light gray areas represent the probability range (from the 5th to the 95th percentile), and the dark gray areas the probability interquartile range (from the 25th to the 75th percentile).

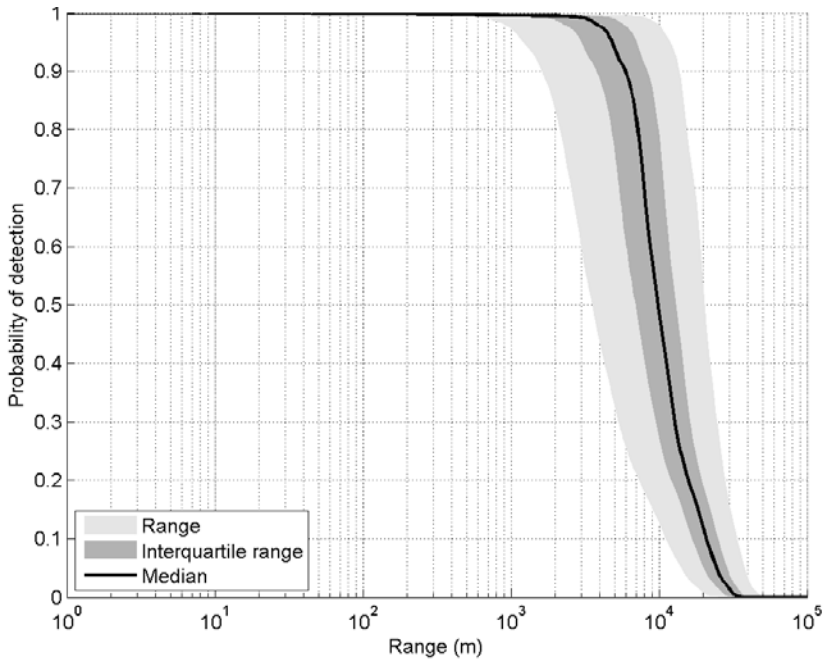

Figure D-12. Detection probability of bowhead moans at Station WN80. The solid black line represents the median probability of detection. The light gray areas represent the probability range (from the 5th to the 95th percentile), and the dark gray areas the probability interquartile range (from the 25th to the 75th percentile).

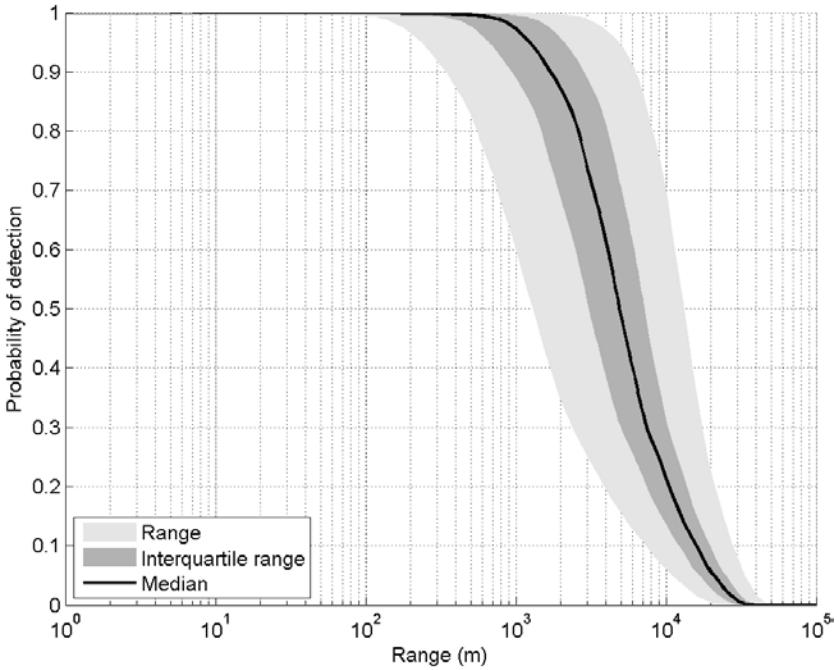

Figure D-13. Detection probability of bowhead moans at Station PL10. The solid black line represents the median probability of detection. The light gray areas represent the probability range (from the 5th to the 95th percentile), and the dark gray areas the probability interquartile range (from the 25th to the 75th percentile).

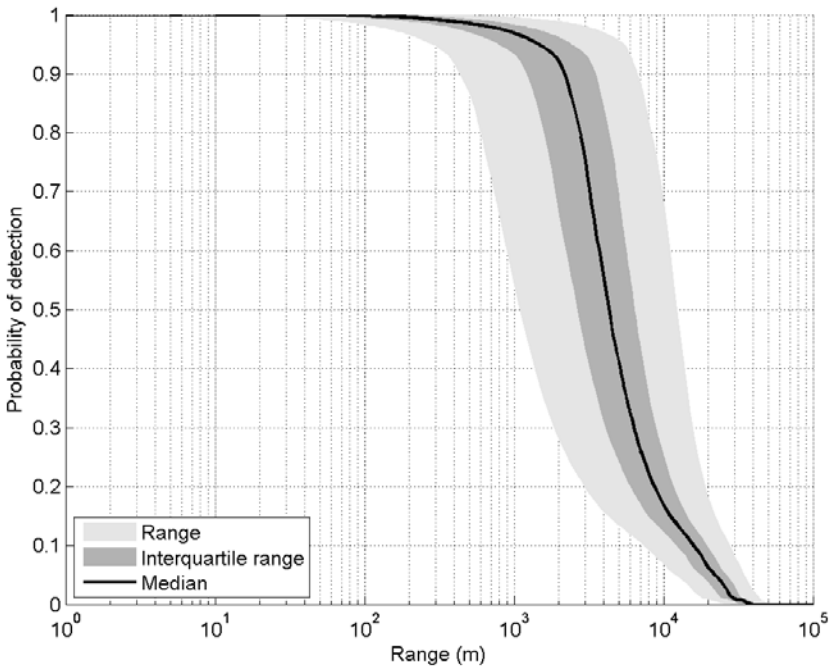

Figure D-14. Detection probability of bowhead moans at Station PL30. The solid black line represents the median probability of detection. The light gray areas represent the probability range (from the 5th to the 95th percentile), and the dark gray areas the probability interquartile range (from the 25th to the 75th percentile).

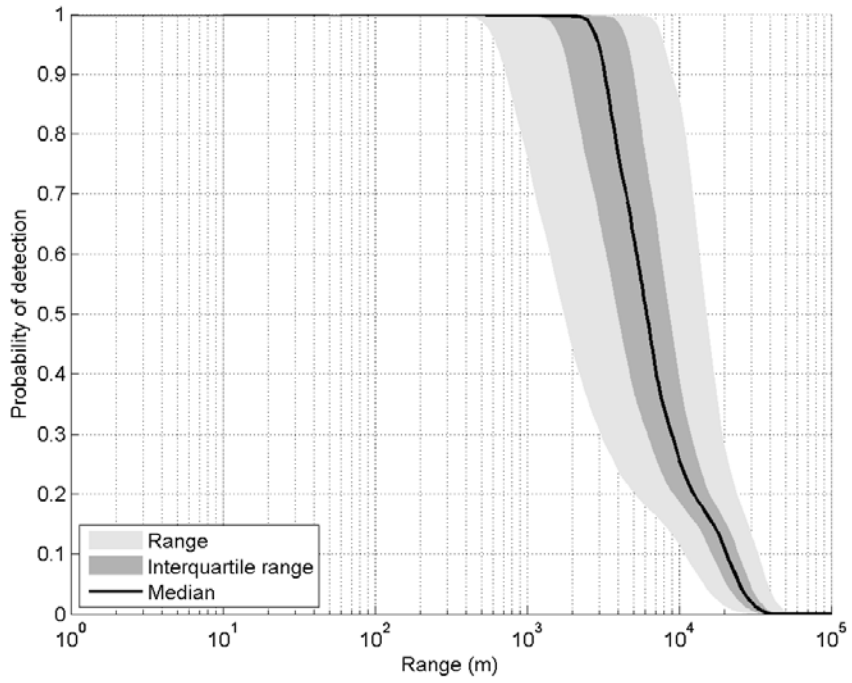

Figure D-15. Detection probability of bowhead moans at Station PL50. The solid black line represents the median probability of detection. The light gray areas represent the probability range (from the 5th to the 95th percentile), and the dark gray areas the probability interquartile range (from the 25th to the 75th percentile).

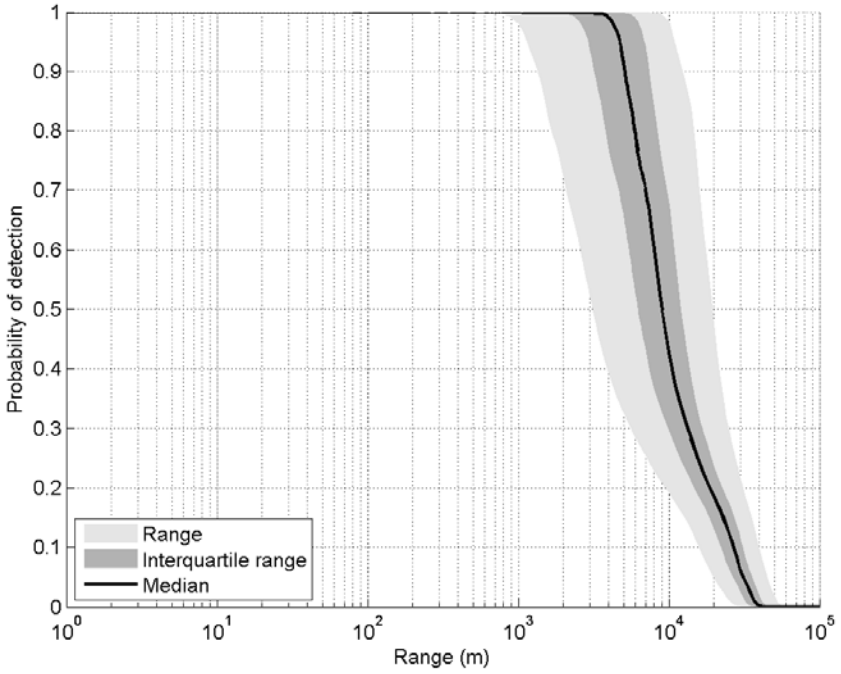

Figure D-16. Detection probability of bowhead moans at Station PLN20. The solid black line represents the median probability of detection. The light gray areas represent the probability range (from the 5th to the 95th percentile), and the dark gray areas the probability interquartile range (from the 25th to the 75th percentile).

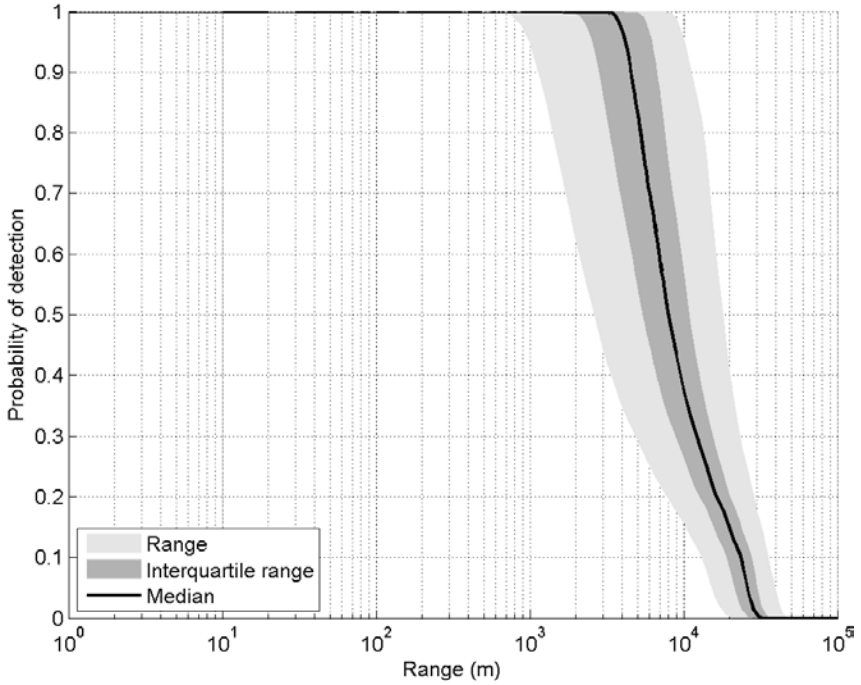

Figure D-17. Detection probability of bowhead moans at Station PLN40. The solid black line represents the median probability of detection. The light gray areas represent the probability range (from the 5th to the 95th percentile), and the dark gray areas the probability interquartile range (from the 25th to the 75th percentile).

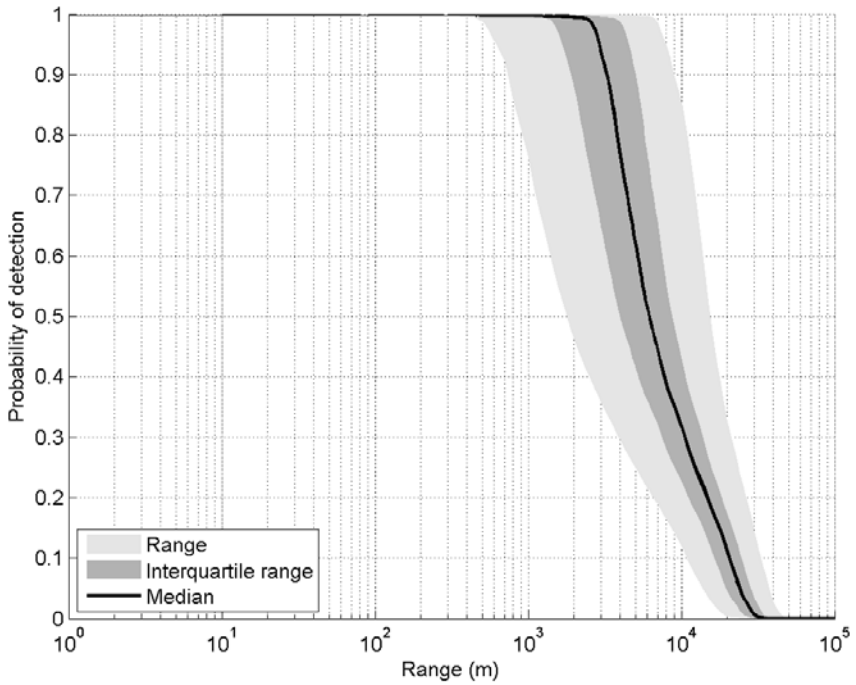

Figure D-18. Detection probability of bowhead moans at Station PLN60. The solid black line represents the median probability of detection. The light gray areas represent the probability range (from the 5th to the 95th percentile), and the dark gray areas the probability interquartile range (from the 25th to the 75th percentile).

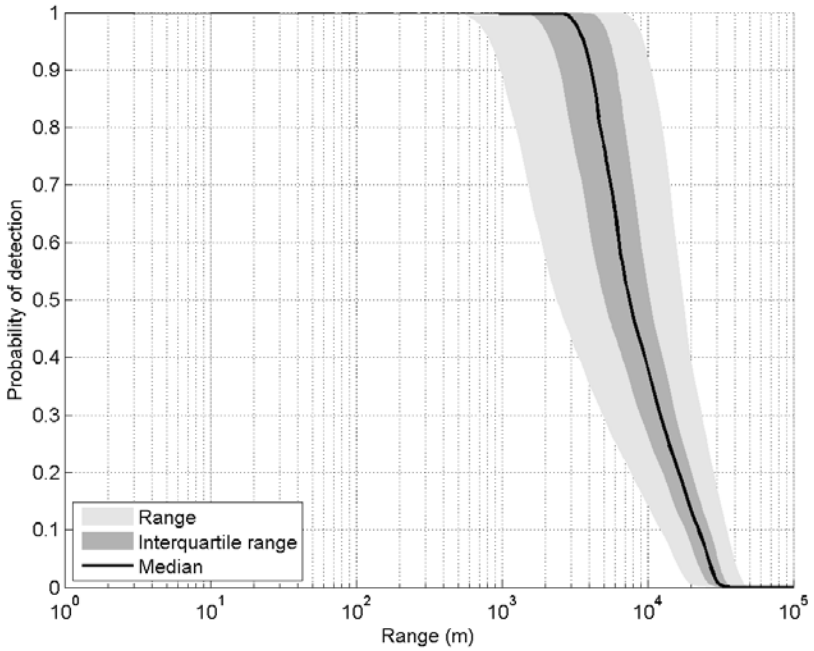

Figure D-19. Detection probability of bowhead moans at Station PLN80. The solid black line represents the median probability of detection. The light gray areas represent the probability range (from the 5th to the 95th percentile), and the dark gray areas the probability interquartile range (from the 25th to the 75th percentile).

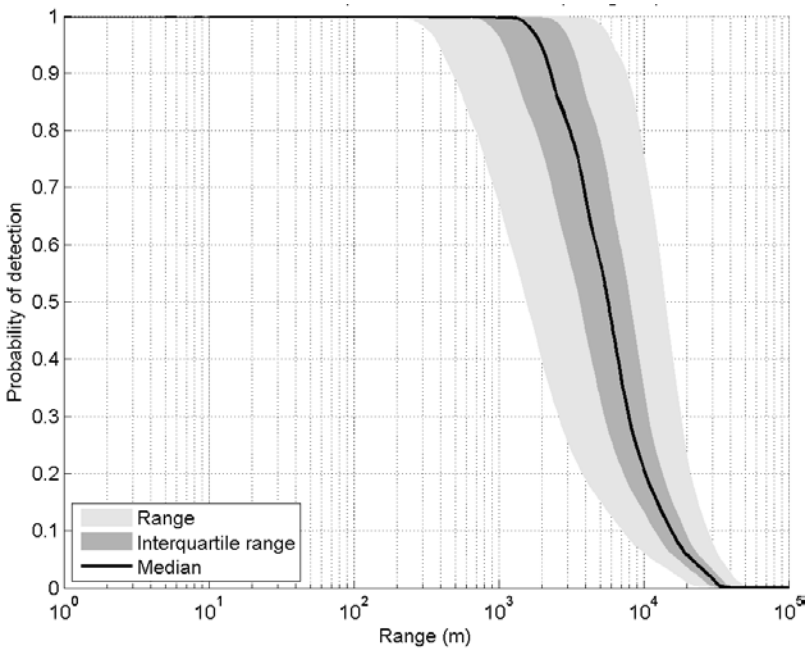

Figure D-20. Detection probability of bowhead moans at Station CL5. The solid black line represents the median probability of detection. The light gray areas represent the probability range (from the 5th to the 95th percentile), and the dark gray areas the probability interquartile range (from the 25th to the 75th percentile).

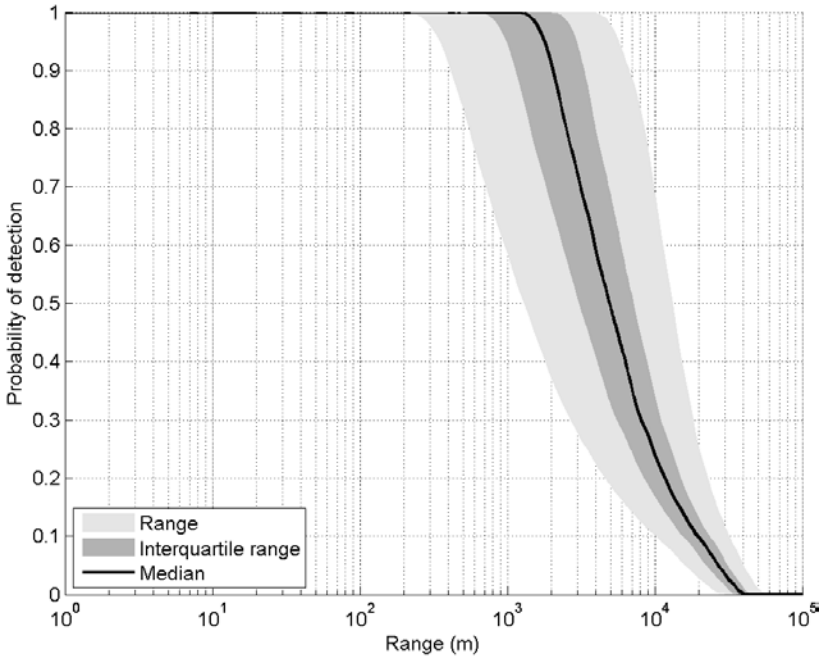

Figure D-21. Detection probability of bowhead moans at Station CL50. The solid black line represents the median probability of detection. The light gray areas represent the probability range (from the 5th to the 95th percentile), and the dark gray areas the probability interquartile range (from the 25th to the 75th percentile).

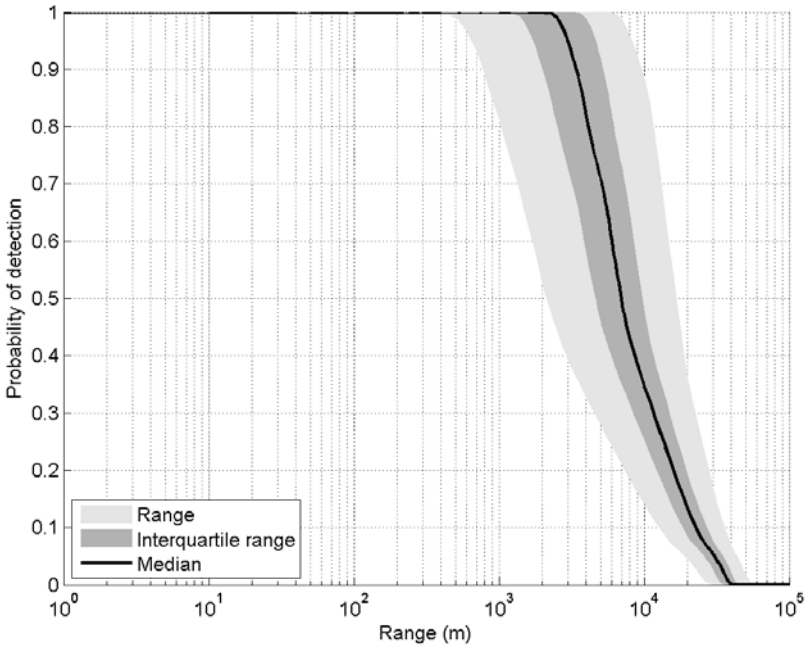

Figure D-22. Detection probability of bowhead moans at Station CLN40. The solid black line represents the median probability of detection. The light gray areas represent the probability range (from the 5th to the 95th percentile), and the dark gray areas the probability interquartile range (from the 25th to the 75th percentile).

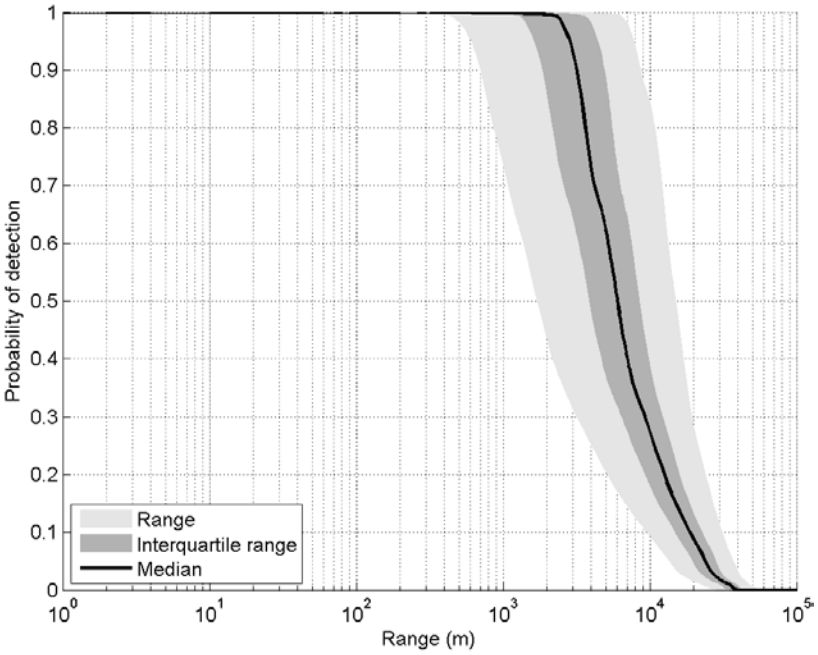

Figure D-23. Detection probability of bowhead moans at Station CLN90B. The solid black line represents the median probability of detection. The light gray areas represent the probability range (from the 5th to the 95th percentile), and the dark gray areas the probability interquartile range (from the 25th to the 75th percentile).

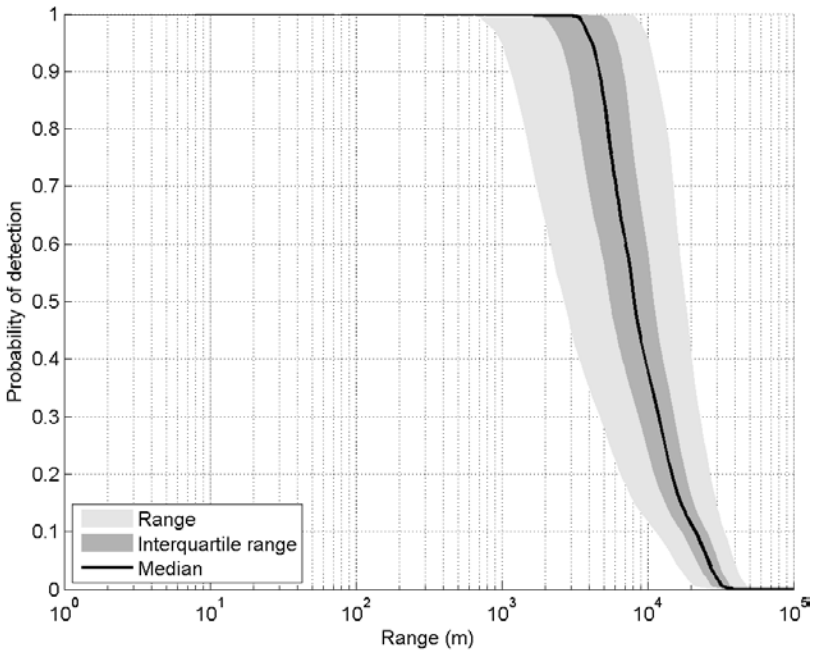

Figure D-24. Detection probability of bowhead moans at Station CLN120B. The solid black line represents the median probability of detection. The light gray areas represent the probability range (from the 5th to the 95th percentile), and the dark gray areas the probability interquartile range (from the 25th to the 75th percentile).

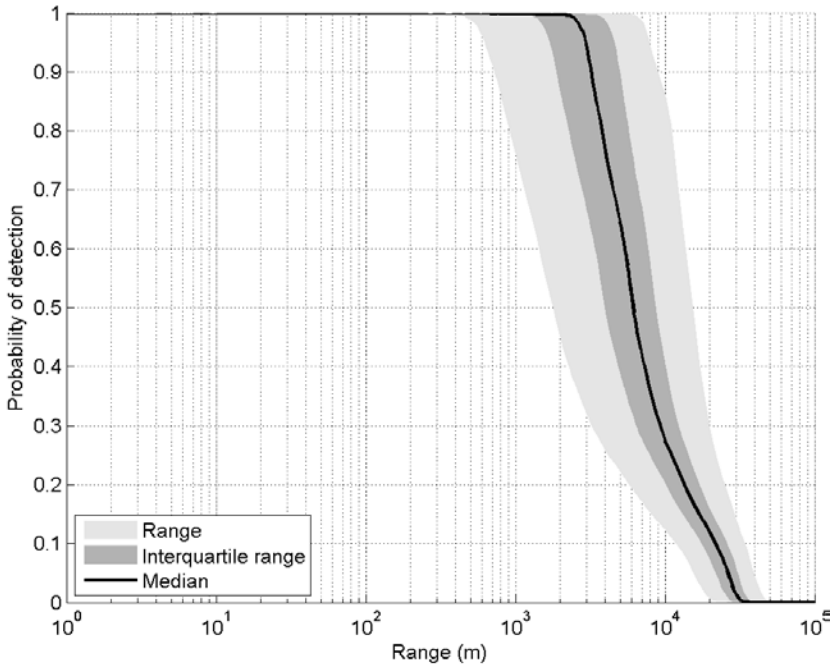

Figure D-25. Detection probability of bowhead moans at Station KL01. The solid black line represents the median probability of detection. The light gray areas represent the probability range (from the 5th to the 95th percentile), and the dark gray areas the probability interquartile range (from the 25th to the 75th percentile).

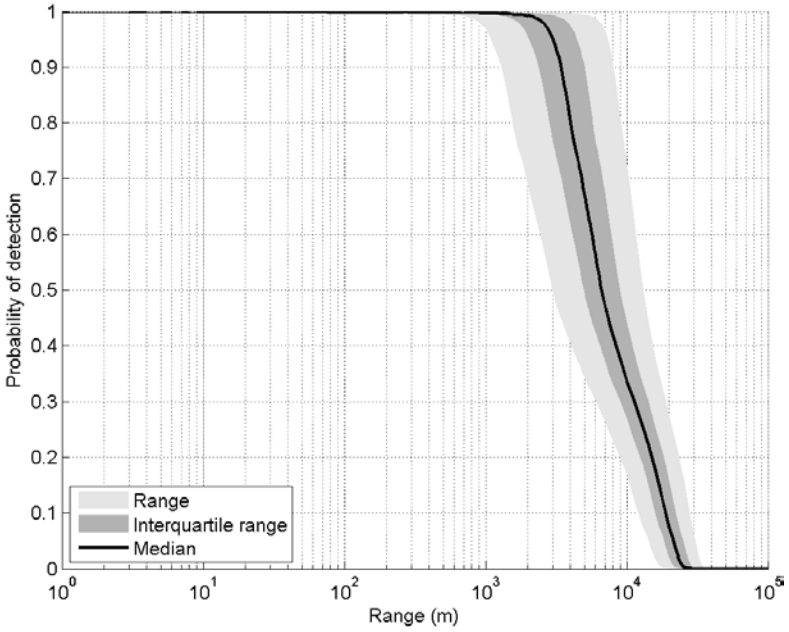

Figure D-26. Detection probability of bowhead moans at Station BGB. The solid black line represents the median probability of detection. The light gray areas represent the probability range (from the 5th to the 95th percentile), and the dark gray areas the probability interquartile range (from the 25th to the 75th percentile).

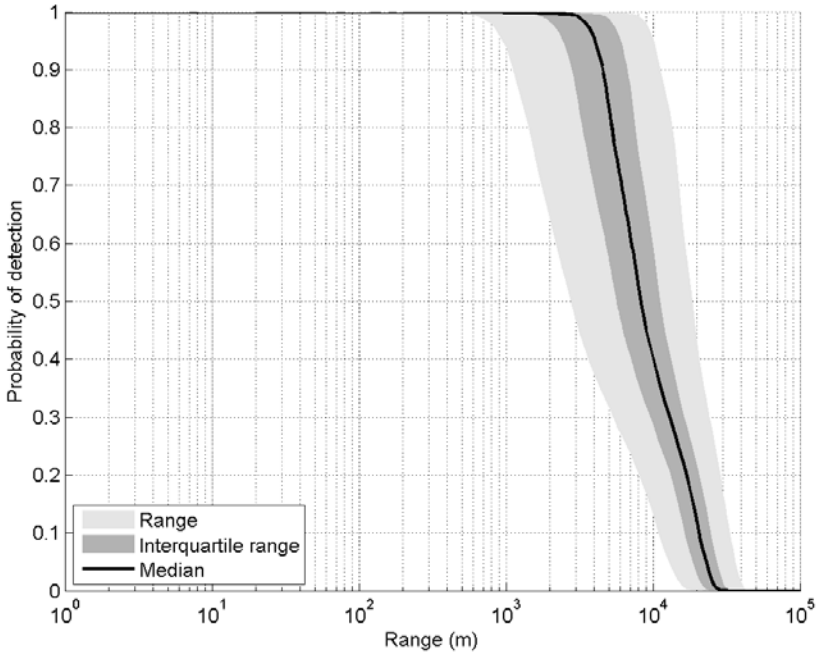

Figure D-27. Detection probability of bowhead moans at Station BGC. The solid black line represents the median probability of detection. The light gray areas represent the probability range (from the 5th to the 95th percentile), and the dark gray areas the probability interquartile range (from the 25th to the 75th percentile).

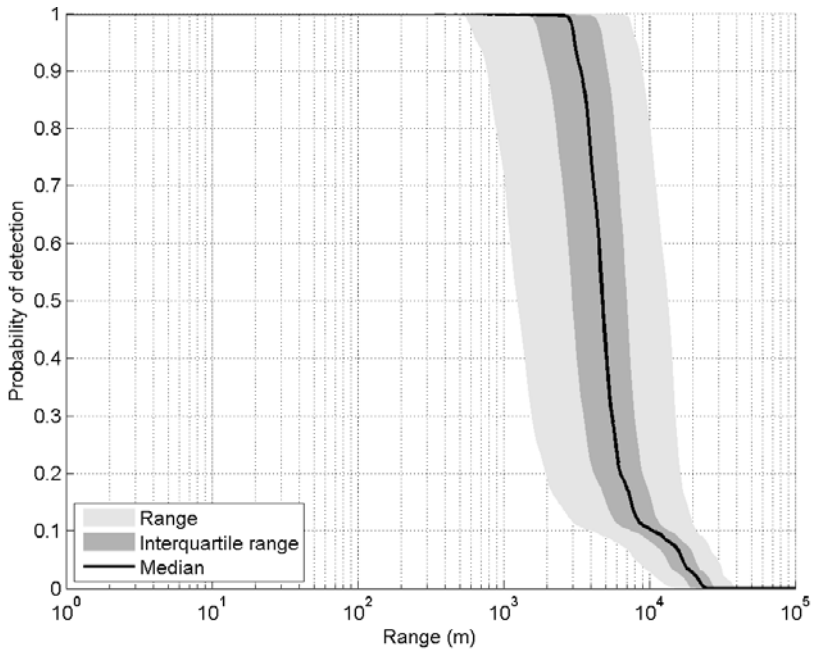

Figure D-28. Detection probability of bowhead moans at Station BGD. The solid black line represents the median probability of detection. The light gray areas represent the probability range (from the 5th to the 95th percentile), and the dark gray areas the probability interquartile range (from the 25th to the 75th percentile).

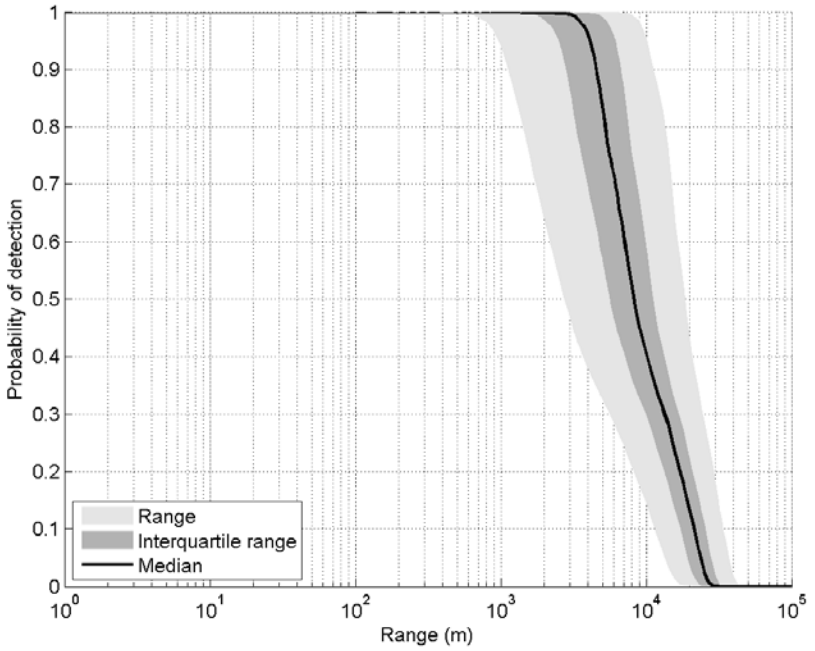

Figure D-29. Detection probability of bowhead moans at Station BGE. The solid black line represents the median probability of detection. The light gray areas represent the probability range (from the 5th to the 95th percentile), and the dark gray areas the probability interquartile range (from the 25th to the 75th percentile).

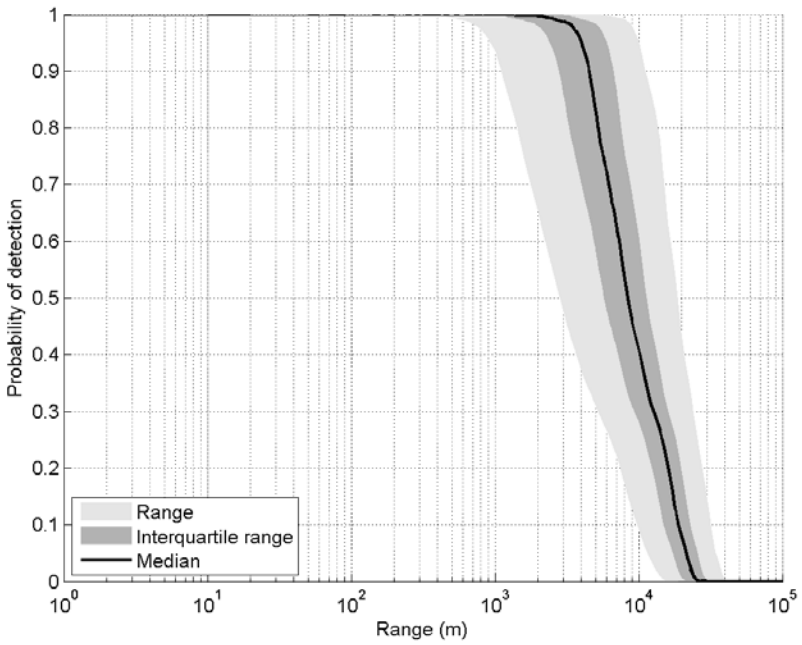

Figure D-30. Detection probability of bowhead moans at Station BGF. The solid black line represents the median probability of detection. The light gray areas represent the probability range (from the 5th to the 95th percentile), and the dark gray areas the probability interquartile range (from the 25th to the 75th percentile).

D.3. Discussion

The maximum detection ranges calculated in this study are consistent with maximum detection ranges reported in the literature. Cummings and Holliday (1985) and Clark et al. (1986) detected bowhead moans off Barrow up to 20 km from a hydrophone, although most of the bowhead moans they localized were less than 10 km away. Figure D-4 and Figure D-5 show similar results with a maximum detection range of approximately 20 km near Barrow.

Figure D-31 shows the median detection probability range at all the Burger stations. The median detection range at BGD is shorter than at other stations because this recorder collected only 10 days of data.

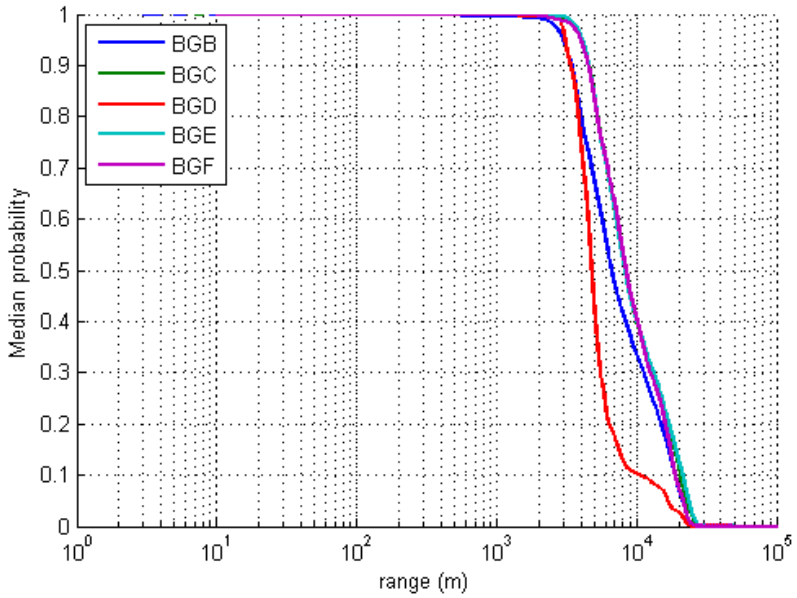

Figure D-31. Comparison of the 50th percentile probability of detection range for bowhead moans at the Burger stations.

Detection range estimates (median probability) calculated in summer 2014 were similar to the ones calculated in summer 2013 (Figure D-32).

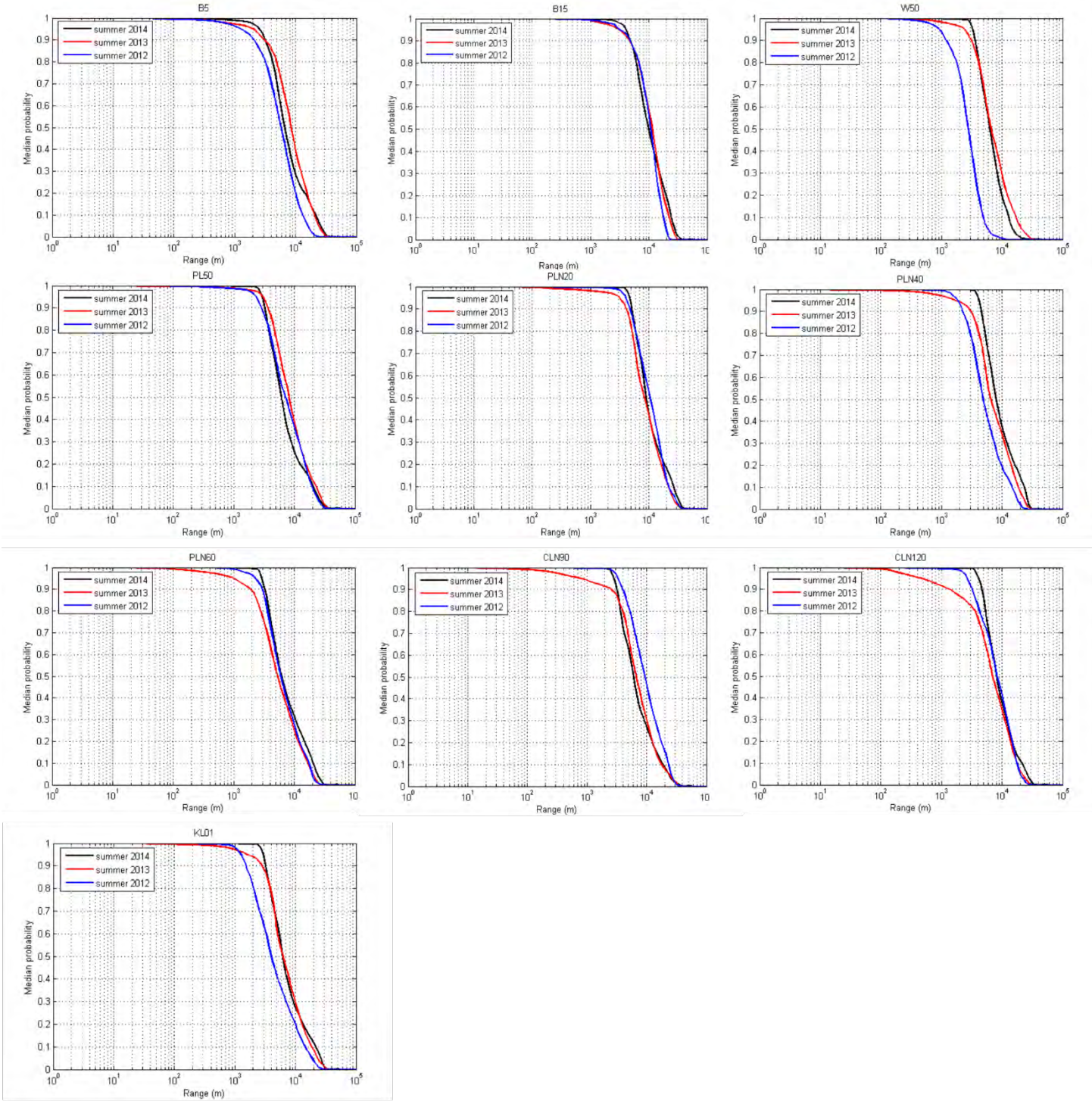

Figure D-32. Comparison of 50th percentile probability of detection range for bowhead moans between summer 2012 through 2014 at several stations.

Blackwell et al. (2012) reported that bowhead moans are directional; sounds are stronger in front of the animal than behind it. Although Equation 1 did not explicitly consider this directionality,

the distribution of source levels used (see Figure D-1) was obtained from bowheads at many different angles, which leads to a measure of average source levels and detectability.

The transmission loss used in this study was calculated in the middle of the water column. This was not optimal because bowheads vocalize at different depths through the water column and transmission loss varies with depth. Although, because the eastern Chukchi Sea is very shallow (~50 m), the variation of received levels due to the variations of source depth is not likely to greatly affect the detection range results.

Appendix E. Interpolation Techniques

There are two main groupings of interpolation techniques used to create surfaces maps from measured points:

- Deterministic: Based on either the extent of similarity or the degree of smoothing of the measured points.
- Geostatistical: Uses the statistical properties of the measured points.

Radial basis function is a deterministic interpolation technique that creates a surface from measured points, based on the degree of smoothing. It calculates predictions from the measured points based on the assumption that the interpolating surfaces should be influenced by a function of their radial distance from a grid point and that the surface must pass through each measured sample value.

There are five different basis functions:

- Thin-plate spline
- Spline with tension
- Completely regularized spline
- Multiquadric function
- Inverse multiquadric function

Each basis function has a different shape and results in a different interpolation surface. Each of the functions has a parameter that controls surface smoothness through a series of elevation samples. Its default value is equal to the average point spacing, assuming the samples are uniformly distributed. The radial basis function used was the inverse multiquadric, given by the equation:

$$B(h) = \frac{1}{\sqrt{h^2+R^2}} \quad (1)$$

where h is the anisotropically scaled distance from the interpolant to the node and R^2 is the kernel parameter that controls surface smoothness. Smoother maps are generated from lower parameter values.

Ordinary Kriging is a geostatistical interpolation technique that relies on both statistical and mathematical methods to create surfaces and assess the uncertainty of the predictions. Ordinary Kriging assumes the model:

$$Z(s) = \mu + \varepsilon(s) \quad (2)$$

where μ is an unknown constant and ε represents errors associated with μ. Ordinary Kriging requires the form and parameter values of the spatial dependence of the spatial process in terms of a semivariogram model. In spatial statistics the theoretical semivariogram is a function (e.g., linear, exponential, Gaussian, and spherical) describing the degree of spatial dependence of a spatial random field or stochastic process. Typically the semivariogram model is not known in advance, and therefore must be estimated, either visually or by an estimation method. The

appropriate model must fit the empirical values by matching the shape of the curve of the experimental variogram to the shape of the curve of the mathematical function.

The main steps in creating a geostatistical model are:

1. Examine the data.
2. Calculate the experimental semivariogram.
3. Fit a theoretical model.
4. Generate the matrices of Kriging equations.
5. Solve the matrices to obtain a predicted value and its associated error for each location in the output surface.

Surface map plots were generated with IDL programming software version 8.2.0, MATLAB version 8.1.0, and ArcGIS version 10.1.

Appendix F. Localization Techniques

F.1. Source Localization

The timespan extracted from each call recording on three synchronized recorders was centered on the detection time, and extended before and after the detection by the maximum sound propagation time between recorders. A 3 × 3 matrix of TDOAs was created based on cross-correlations between each pair or recorders. Each row in the matrix represented the computed TDOAs between one recorder and the two others. Diagonal values were set to zero because they correspond to a recording's delay relative to itself. The strength of each matrix row was scored by summing the magnitude of its three strongest correlations. TDOAs from the row with the highest score were used to calculate the localization using its three strongest correlations. Some cross-correlations were weak because the signal to noise ratio (SNR) was too low or a bowhead vocalization was not recorded by enough recorders. In some cases, no localization solution could be found for a detection group and that group was discarded.

The linear equation approach describes the algebraic relation between the TDOA and the locations of the source and the receivers. Defining one of the receivers as the origin, the source location(s) from a three-receiver array is obtained as:

$$s = \frac{1}{2} R^{-1} b - c^2 \delta \tau_1 R^{-1} \tag{1}$$

where c is sound velocity, δ is the TDOA vector, $\delta = \left[\delta_{12}, \delta_{13}\right]^T$, b is given by $b_i = r_{x(i)}^2 + r_{y(i)}^2 - c^2 \delta_{1(i)}^2$, τ_1 is time of arrival from source to receiver reference, and R represents the receiver matrix:

$$R = \begin{bmatrix} r_{x(2)} & r_{y(2)} \\ r_{x(3)} & r_{y(3)} \end{bmatrix} \tag{2}$$

Solving for τ_1:

$$\tau_1 = \frac{c a_2 \pm \sqrt{c^2 a_2^2 - \left(c^2 a_3 - 1\right) a_1}}{2c\left(c^2 a_3 - 1\right)} \tag{3}$$

where $a_2 = \left(R^{-1}\tau\right)^T \left(R^{-1} b\right)$ and $a_3 = \left(R^{-1}\tau\right)^T \left(R^{-1}\tau\right)$.

Substituting Equation 2 into Equation 1, the source location s is obtained. Two positive solutions correspond to two possible source positions. Negative and complex solutions are discarded, as they have no physical solution or meaning (Wahlbergh et al. 2001, Vallarta 2009).

References

[NOAA] National Oceanic and Atmospheric Administration. 2008. IMS daily Northern Hemisphere snow and ice analysis at 4 km and 24 km resolution. National Snow and Ice Data Center. http://dx.doi.org/10.7265/N52R3PMC

Abbot, T.A., V.E. Premus, and P.A. Abbot. 2010. A real-time method for autonomous passive acoustic detection-classification of humpback whales. *Journal of the Acoustical Society of America* 127(5): 2894-2903. http://dx.doi.org/10.1121/1.3365255.

Ainslie, M.A. and J.G. McColm. 1998. A simplified formula for viscous and chemical absorption in sea water. *Journal of the Acoustical Society of America* 103(3): 1671-1672.

Arveson, P.T. and D.J. Vendittis. 2000. Radiated noise characteristics of a modern cargo ship. *Journal of the Acoustical Society of America* 107(1): 118-129.

Austin, M. 2012. *Modelling three-dimensional sound propagation in wedge environments*. PhD Thesis. University of Victoria, Victoria.

Balanda, K.P. and H.L. MacGillivray. 1988. Kurtosis: A critical review. *American Statistician* 42: 111.

Blackwell, S.B., T.L. McDonald, K.H. Kim, L.A.M. Aerts, W.J. Richardson, J. Greene, Charles R., and B. Streever. 2012. Directionality of bowhead whale calls measured with multiple sensors. *Marine Mammal Science* 28(1): 200-212. http://dx.doi.org/10.1111/j.1748-7692.2011.00464.x.

Breiman, L., J. Friedman, R. Olshen, and C. Stone. 1984. *Classification and regression trees*. Wadsworth International Group, Belmont, California. 358.

Breiman, L. 2001. Random Forests. *Machine Learning* 45: 5-32.

Clark, C.W., W.T. Ellison, and K. Beeman. 1986. *An acoustic study of bowhead whales, Balaena mysticetus, off Point Barrow, Alaska during the 1984 spring migration*. Report to the North Slope Borough, Department of Wildlife Management. Marine Acoustics, Barrow, Alaska. 145 pp. http://books.google.ca/books?id=49dFAAAAYAAJ.

Clemins, P.J., M.T. Johnson, K.M. Leong, and A. Savage. 2005. Automatic classification and speaker identification of African elephant (*Loxodonta africana*) vocalizations. *Journal of the Acoustical Society of America* 117(2): 956-963.

Clemins, P.J. and M.T. Johnson. 2006. Generalized perceptual linear prediction features for animal vocalization analysis. *Journal of the Acoustical Society of America* 120(1): 527-534.

Cummings, W.C. and D.V. Holliday. 1985. Passive acoustic location of bowhead whales in a population census off Point Barrow, Alaska. *Journal of the Acoustical Society of America* 78(4): 1163-1169.

Cummings, W.C. and D.V. Holliday. 1987. Sounds and source levels from bowhead whales off Pt. Barrow, Alaska. *Journal of the Acoustical Society of America* 82(3): 814-821. http://scitation.aip.org/content/asa/journal/jasa/82/3/10.1121/1.395279.

Davis, J. and M. Goadrich. 2006. *The relationship between Precision-Recall and ROC curves*. *Proceedings of the 23rd international conference on machine learning*. ACM, Pittsburgh, PA, pp. 233-240. www.autonlab.org/icml_documents/.../030_The_Relationship_Bet.pdf.

Delarue, J., M. Laurinolli, and B. Martin. 2009. Bowhead whale (*Balaena mysticetus*) songs in the Chukchi Sea between October 2007 and May 2008. *Journal of the Acoustical Society of America* 126(6): 3319-3328. http://link.aip.org/link/?JAS/126/3319/1.

Delarue, J., B. Martin, X. Mouy, J. MacDonnell, J. Vallarta, and N. Chorney. 2011. *Northeastern Chukchi Sea Joint Acoustic Monitoring Program 2009-2010*. *In*: Hannay, D. (ed.). Technical report for ConocoPhillips Company, Shell Exploration & Production Company, and Statoil USA E&P, Inc. by JASCO Applied Sciences, Dartmouth, NS. 156 pp.

https://www.chukchiscience.com/Portals/0/Public/Science/AcousticMoorings/2010_CSESP_Acoustics_Final_Report.pdf.

Fristrup, K.M. and W.A. Watkins. 1993. *Marine animal sound classification*. Woods Hole Oceanographic Institute Technical Report WHOI-94-13. Woods Hole Oceanographic Institution. 32 pp. http://darchive.mblwhoilibrary.org:8080/bitstream/handle/1912/546/WHOI-94-13.pdf?sequence=1.

Gillespie, D. 2004. Detection and classification of right whale calls using an 'edge' detector operating on a smoothed spectrogram. *Canadian Acoustics* 32(2): 39-47.

Hannay, D. and R. Racca. 2005. *Acoustic Model Validation*. Document Number 0000-S-90-04-T-7006-00-E, Revision 02. Technical report for Sakhalin Energy Investment Company Ltd. by JASCO Research Ltd. 34 pp.

Karlsen, J., A. Bisther, C. Lydersen, T. Haug, and K. Kovacs. 2002. Summer vocalisations of adult male white whales (*Delphinapterus leucas*) in Svalbard, Norway. *Polar Biology* 25(11): 808-817. http://link.springer.com/article/10.1007/s00300-002-0415-6.

Kogan, J.A. and D. Margoliash. 1998. Automated recognition of bird song elements from continuous recordings using dynamic time warping and hidden Markov models: A comparative study. *Journal of the Acoustical Society of America* 103(4): 2185-2196. http://scitation.aip.org/content/asa/journal/jasa/103/4/10.1121/1.421364.

MacDonnell, J. and B. Martin. 2011. Estimating bowhead whale communications space using measured and modeled data. *Journal of the Acoustical Society of America* 129(4): 2574-2574. http://scitation.aip.org/content/asa/journal/jasa/129/4/10.1121/1.3588494.

MacDonnell, J., J. Vallarta, and B. Martin. 2011. Source Levels of Bowhead Moans from Acoustic Measurements and Propagation Modeling in the Chukchi Sea. *The Alaska Marine Science Symposium*. Jan 17-21 2011, Anchorage, Alaska.

Mellinger, D.K. and C.W. Clark. 1997. Methods for automatic detection of mysticete sounds. *Marine and Freshwater Behaviour and Physiology* 29(1-4): 163-181. http://dx.doi.org/10.1080/10236249709379005.

Mellinger, D.K. and C.W. Clark. 2000. Recognizing transient low-frequency whale sounds by spectrogram correlation. *Journal of the Acoustical Society of America* 107(6): 3518-3529. http://scitation.aip.org/content/asa/journal/jasa/107/6/10.1121/1.429434.

Mellinger, D.K. 2001. *Ishmael 1.0 User's Guide*. Technical Memorandum OAR PMEL-120. NOAA. http://www.pmel.noaa.gov/pubs/PDF/mell2434/mell2434.pdf.

Mellinger, D.K. 2004. A comparison of methods for detecting right whale calls. *Canadian Acoustics* 32(2): 55-65. http://jcaa.caa-aca.ca/index.php/jcaa/article/view/1588.

Mellinger, D.K. and C.W. Clark. 2006. MobySound: A reference archive for studying automatic recognition of marine mammal sounds. *Applied Acoustics* 67(11-12): 1226-1242.

Mellinger, D.K. and J.W. Bradbury. 2007. *Acoustic measurement of marine mammal sounds in noisy environments*. *Proceedings of the Second International Conference on Underwater Acoustic Measurements: Technologies and Results*, Heraklion, Greece, pp. 273-280. ftp://ftp.pmel.noaa.gov/newport/mellinger/papers/Mellinger+Bradbury07-BioacousticMeasurementInNoise-UAM,Crete.pdf.

Mellinger, D.K., K.M. Stafford, S.E. Moore, R.P. Dziak, and H. Matsumoto. 2007. An overview of fixed passive acoustic observation methods for cetaceans. *Oceanography* 20(4): 36–45.

Mellinger, D.K., S.W. Martin, R.P. Morrissey, L. Thomas, and J.J. Yosco. 2011. A method for detecting whistles, moans, and other frequency contour sounds. *Journal of the Acoustical Society of America* 129(6): 4055-4061.

Mouy, X., D. Leary, B. Martin, and M. Laurinolli. 2008. *A comparison of methods for the automatic classification of marine mammal vocalizations in the Arctic*. *New Trends for Environmental*

Monitoring Using Passive Systems, 2008, 14-17 Oct. Institute of Electrical and Electronics Engineers, pp. 1-6.

Mouy, X., M. Bahoura, and Y. Simard. 2009. Automatic recognition of fin and blue whale calls for real-time monitoring in the St. Lawrence. *Journal of the Acoustical Society of America* 126(6): 2918-2928. http://scitation.aip.org/content/asa/journal/jasa/126/6/10.1121/1.3257588.

Nosal, E. 2008. Flood-fill algorithms used for passive acoustic detection and tracking. *Proc. IEEE Workshop and Exhibition on New Trends for Environmental Monitoring using Passive Systems*. 14-17 Oct. 2008. IEEE Oceanic Engineering Society, Hyeres, France. 1-5 pp. http://www.soest.hawaii.edu/ore/faculty/nosal/publications/ConferencePapers/2008_Nosal_Passive.pdf.

Oswald, J.N., J. Barlow, and T.F. Norris. 2003. Acoustic identification of nine delphinid species in the eastern tropical Pacific Ocean. *Marine Mammal Science* 19(1): 20-37. http://swfsc.noaa.gov/uploadedFiles/Divisions/PRD/Programs/Coastal_Marine_Mammal/Oswald.pdf.

Oswald, J.N., S. Rankin, J. Barlow, and M.O. Lammers. 2007. A tool for real-time acoustic species identification of delphinid whistles. *Journal of the Acoustical Society of America* 122(1): 587-595. http://scitation.aip.org/content/asa/journal/jasa/122/1/10.1121/1.2743157.

Rankin, S. and J. Barlow. 2005. Source of the North Pacific “boing” sound attributed to minke whales. *Journal of the Acoustical Society of America* 118(5): 3346-3351.

Stafford, K.M. 1995. *Characterization of Blue Whale Calls from the Northeast Pacific and Development of a Matched Filter to Locate Blue Whales on U.S. Navy SOSUS (SOund SUrveillance System) arrays*. M.Sc. Thesis. Oregon State University, Corvallis, OR. 89 pp. http://ir.library.oregonstate.edu/xmlui/bitstream/handle/1957/15875/StaffordKathleenMary1995.pdf?sequence=1.

Struzinski, W.A. and E.D. Lowe. 1984. A performance comparison of four noise background normalization schemes proposed for signal detection systems. *Journal of the Acoustical Society of America* 76(6): 1738-1742. http://scitation.aip.org/content/asa/journal/jasa/76/6/10.1121/1.391621.

Thompson, P.O., W.C. Cummings, and S.J. Ha. 1986. Sounds, source levels, and associated behavior of humpback whales, southeast Alaska. *Journal of the Acoustical Society of America* 80(3): 735-740.

Urick, R.J. 1983. *Principles of Underwater Sound*. 3rd edition. McGraw-Hill, New York, London. 423.

Vallarta, J. 2009. *The Significance of Passive Acoustic Array-Configuration on Sperm whale Range Estimation when using the Hyperbolic Algorithm*. PhD Thesis. Heriot-Watt University, Edinburgh, UK. 225 pp.

Wahlbergh, M., B. Mohl, and P.T. Madsen. 2001. Estimating source position accuracy of a large-aperture hydrophone array for bioacoustics. *Journal of the Acoustical Society of America* 109: 397-406.

Watkins, W.A. and G.C. Ray. 1977. Underwater sounds from ribbon seal, *Phoca (Histriophoca) fasciata*. *Fishery Bulletin* 75: 450-453. https://darchive.mblwhoilibrary.org/bitstream/handle/1912/6185/Watkins_Ray_1977%20ribbon%20seal%20vocalizations.pdf?sequence=1.