Suggested abbreviated title: Single-sensor density estimation of broadband clicks

Single-sensor, cue-counting population density estimation: Average probability of detection of broadband clicks.

Elizabeth T. Küsel and Martin Siderius

Northwest Electromagnetics and Acoustics Research Laboratory,

Portland State University,

1900 SW 4th Avenue,

Portland,

OR 97201

David K. Mellinger

Cooperative Institute for Marine Resources Studies,

Oregon State University,

2030 Marine Science Drive,

Newport,

OR 97365

(Dated: October 23, 2015)

Abstract

Odontocete echolocation clicks have been used as a preferred cue for density estimation studies from single-sensor data sets, which require estimating detection probability as a function of range. Many such clicks can be very broadband in nature, with 10-dB bandwidths of 20 to 40 kHz or more. Detection distances are not readily obtained from single-sensor data. Therefore, the average detection probability is estimated in a Monte Carlo simulation using the passive sonar equation along with transmission loss calculations to estimate the signal-to-noise ratio (SNR) of tens of thousands of click realizations. Continuous-wave (CW) analysis, i.e., single-frequency analysis, is inherent to basic forms of the passive sonar equation. Using CW analysis with the click's center frequency while disregarding its bandwidth has been shown to introduce bias to detection probabilities and hence to population estimates. In this study, the effect of highly broadband clicks on density estimates is further examined. The usage of transmission loss as an appropriate measure for calculating click SNR is also discussed. The main contributions from this research are: 1. an alternative approach to estimate average probability of detection of broadband clicks, and 2. the effects of multipath clicks on population density estimates.

PACS numbers: 43.30.Sf,43.80.Ka

I. INTRODUCTION

The main objective of this work is to further investigate the methodology that has been used to estimate detection probabilities of highly broadband clicks recorded by single, fixed instruments and used in cetacean population density estimates. The development of passive acoustic-based density estimation methods for marine mammal studies in general, has been an active area of research. Two recent review papers on passive acoustic techniques for animal population density estimation (Thomas and Marques, 2012; Marques *et al.*, 2013) summarize the different approaches that have been used and the challenges that they present.

In the single, fixed sensor modality, which is the scope of this work, a cue-counting approach which scales the number of detected calls by their probability of detection and call production rate is used (Küsel *et al.*, 2011). This methodology works best for those species which produce easily detectable and distinguishable calls (Marques *et al.*, 2011), as is, in general, the case of odontocete echolocation clicks. The cue-counting density estimation formula presented by Marques *et al.* (2009) and used by Küsel *et al.* (2011) for the single-sensor case is given by,

$$\hat{D} = \frac{n_c(1-\hat{c})}{\pi w^2 \hat{P} T \hat{r}}. \tag{1}$$

In Eq. 1, n_c is the number of auto-detected clicks in a given time period T (in seconds), $\hat{c}$ is the estimated proportion of false positive detections determined by a human analyst who examines some percentage of the auto-detected clicks, w (in meters) is the distance beyond which no cues are expected to be detected, or in other words, an assumed detection range, the cue rate $\hat{r}$ (cues/s) is a measure of how often cues are produced by an animal and converts the total number of detected cues into the number of animals it represents, and $\hat{P}$ is the estimated average probability of detecting a cue within distance w. Finally, $\hat{D}$ is the estimated density commonly given as the number of animals per 1000 km^2. Parameters with a *hat* correspond to estimated quantities since they are not known with certainty. While different detection and classification techniques, as well as whale call production rates are important to passive acoustic density estimation, they are beyond the objectives of this paper and are assumed known here.

The biggest challenge in using Eq. 1 is to correctly estimate the average probability of detecting cues, $\hat{P}$, where cues will be referred henceforth as clicks. Detection distances are not readily realizable from single-sensor data sets. Therefore, $\hat{P}$ is assumed as a function of signal-to-noise ratio (SNR) measured from clicks manually detected by a human analyst. The measured SNRs are compared to clicks detected by a computer algorithm to construct a probability curve as a function of click SNR, or the detection function. Thousands of SNRs of simulated received signals are then compared against the detection function and an average probability of detection is estimated. The simulated SNRs are computed by using the basic form of the passive sonar equation, which is inherently a continuous-wave (CW) analysis, i.e., a single-frequency analysis. Calculation of SNR by means of the sonar equation depends not only on the characteristics of the source, but also on environmental properties between source and receiver. Most importantly, transmission loss (*TL*) is usually calculated at the click's center frequency by using either a simple spherical spreading law plus absorption or a complex acoustic propagation model. However, cetacean echolocation clicks are high-frequency and highly broadband signals, with 10-dB bandwidths of 20 to 40 kHz or more. So, what is the correct way to model such whale calls?

Recently, Ainslie (2013) showed through analytical formulations that considering transmission loss by using CW analysis with the clicks center frequency while disregarding its bandwidth introduces bias to detection probabilities and hence to population density estimates. In order to correctly apply the sonar equation to estimate detection probabilities, Ainslie (2013) proposed the use of a broadband propagation factor correction, which is based on the frequency dependence of absorption on propagation losses. Such correction is derived from a top hat function and assumes spherical spreading and linear dependence on frequency of the absorption coefficient (Ainslie, 2010). It has previously been used by von Benda-Beckmann *et al.* (2010) to estimate detection probabilities of groups of diving Blainville's beaked whales in the Tongue of the Ocean, Bahamas, with good agreement between measured and modeled detection functions.

In the present work a different approach than the one suggested by Ainslie (2013) is investigated. Using simple modeling experiments, the bias in the sonar equation estimates of detection probability and its effect on density estimates is computed. Because synthetic data sets provide ground truth against which to test the methodology, the effects of density estimations using the click's full bandwidth compared to calculations done using only its center frequency are assessed. In addition, the usage of transmission loss as an appropriate measure for calculating the SNR of received clicks is discussed. An alternative approach to estimating the average probability of detection of broadband clicks based on the calculation of ray arrivals is also presented. Last, the issue of including multipath clicks in the analysis is considered.

The remainder of the paper is organized as follows. Section II talks about a data set recorded off the Kona coast of Hawai'i containing highly-broadband false killer whale (*Pseudorca crassidens*) echolocation clicks and gives the motivation for the analysis that follows. Section III describes a simple computational example that illustrates how different frequencies are related to different detection ranges. The creation and analysis of a synthetic data set containing broadband clicks resembling those of false killer whales are presented in Sec. IV.A and IV.B. Discussion on proper ways to treat broadband calls as well as the usage of the parameter *TL* and the sonar equation ensue (Sec. IV.C-E). Section V discusses the main points learned about population density estimation from broadband calls recorded at single sensors, and draws conclusions for future work.

II. HAWAI'I 2010 DATA SET

A broadband acoustic data set, with sampling frequency of 200 kHz, recorded by an autonomous bottom-moored, High-frequency Acoustic Recording Package, or HARP (Wiggins and Hildebrand, 2007), was available to this project through collaboration with Dr. Erin Oleson at the Pacific Islands Fisheries Science Center, HI. The HARP was deployed close to the ocean bottom at a depth of 620 m off the Kona coast of the Big Island of Hawai'i.

The data were recorded in 2010 when the instrument was deployed for a total of 181 days.

Preliminary analysis of the data set included plotting long-term spectral averages to identify periods with false killer whale clicking activity, and using a click detection algorithm described in Soldevilla *et al.* (2008) and Roch *et al.* (2011). Results of this preliminary analysis showed that only 16 out of the 181 days the HARP was deployed had detections of false killer whale clicks. Moreover, only a few hours of clicking activity was observed in those few days. For the current analysis a random period of just over two hours of false killer whale activity recorded on May 22nd (continuous recording) was selected.

A first striking observation when looking at spectrograms of the data set is the highly broadband, high-frequency nature of false killer whale echolocation clicks. Such features, illustrated by 2-second long spectrograms of the data and shown in Figs. 1 and 2, raised questions regarding the single-sensor density estimation methodology being used. Looking at the spectrogram in Fig. 1 for example, it is observed that while click bandwidth seems to be roughly between 20 and 60 kHz in the first 1-second of the data snapshot, in the remaining second clicks are observed to span frequencies beyond 90 kHz.

According to the single-sensor methodology, clicks that are manually annotated on a spectrogram by a human analyst are compared to clicks that are automatically detected using a species characteristic band, say from 20 to 60 kHz, and finally also compared to simulated clicks whose SNR were computed at just a single frequency. In order for the simulated SNRs to be consistent with what is measured from the data set the click bandwidth needs to be taken into account. Higher frequencies experience greater attenuation in the ocean. Simply considering the center frequency for the calculation of SNRs would cause a bias in detection distances. So, what is the best way to simulate such broadband signals? More importantly, how should a specific band be chosen such that detection distances are not biased?

Further examination of Fig. 1, more specifically of the time series data, shows potential multipath arrivals as the shorter impulse signals succeeding the high amplitude ones. Multipath is often hard to discern and separate from first arrivals and clicks from other animals, especially in single-sensor data sets. It is also noted that the parameter TL used in the sonar equation is in fact the sum of all ray arrivals, or multipaths. This implies that an SNR simulated with the sonar equation actually corresponds to a sum of SNRs from all contributing ray paths, instead of being from only one arrival, direct or reflected for example.

Finally, a good amount of diffuse reverberation was also observed in the data set, and is illustrated in Fig. 2. Reverberation appears in the spectrogram as energy smeared over a longer time interval than the click's impulsive duration. Even though click onset can be distinguishable, a clear end or duration can not be assigned to them. Moreover, commonly used propagation models do not account for reverberation phenomena. Last, due to the way HARP data are recorded, many clicks were also clipped, which is observed on the spectrogram by signals spanning the entire spectral band.

III. A SIMPLE DENSITY ESTIMATION EXAMPLE

A broadband signal (e.g. an echolocation click), spanning the 20 to 60 kHz frequency range is assumed. When working with data recorded by a single sensor, the distances at which detected sounds were produced can not be easily realizable. Therefore, computational methods are used to estimate the probability of detection of a click as a function of distance from the sensor. Most of these methods are based on the analysis of a single frequency, usually the center frequency of the signal of interest. Using a single frequency introduces bias in detection ranges, and hence in the probability of detection.

One simple way to illustrate the issue of using different frequencies in the computations is shown in Fig. 3. Received level, i.e. source level (SL) (assumed to be 155 dB re 1 μPa^2/Hz) minus transmission loss (TL), was plotted as a function of range considering the single frequencies of 20 and 40 kHz. Corresponding hypothetical noise levels (NL) at the two frequencies considered are also plotted (straight horizontal lines). A detection occurs, assuming for example a detection threshold of 0 dB for illustration purposes, when the received level is higher than the back- ground noise level. The maximum detection ranges are indicated by the vertical lines in Fig. 3, at approximately 2.5 and 5 km for the 40 and

20 kHz frequencies, respectively. In this illustration transmission loss was calculated using a simple spherical spreading law plus absorption due to high frequencies (Urick, 1983),

$$TL = 20\log_{10}(r) + \alpha(r/1000). \tag{2}$$

In the equation above, r is the horizontal distance between animal and recording sensor in meters, and α is the frequency-dependent attenuation in dB/km calculated by the formula (Jensen *et al.*, 2011),

$$\alpha = 3.3 \times 10^{-3} + \frac{0.11 f^2}{1 + f^2} + \frac{44 f^2}{4100 + f^2} + 3 \times 10^{-4} f^2, \tag{3}$$

where f is frequency in kHz.

As observed in Fig. 3, detection distances can vary significantly depending on the choice of frequency for the calculation of transmission loss. Considering the broadband nature of marine mammal echolocation clicks, higher frequencies will be attenuated faster and, given the right propagation conditions, will only be detected when animals are at closer distances to the recording instrument.

The simple computational experiment described in this section was devised to show the influence of choosing two different single frequencies to calculate transmission loss plus attenuation, following Eqs. 2 and 3, from which the probability of detection is estimated. For such purpose, it was assumed that all detections were made at 20 kHz but the average probability of detection, $\hat{P}$, was estimated at the center frequency of the click, at 40 kHz. This simple example was based on Eq. 1 and the single-sensor density estimation methodology, but introducing a few assumptions as detailed in the next subsection.

A. The Experiment

To create a simple synthetic data, it was assumed that 1000 animals were uniformly distributed across a circular area of ocean with radius R equal to 20 km. A recording hydrophone is assumed to be located at the center of this area. It is noted that such information on the true number of animals and the area where they are distributed is never known in practice. The 1000 synthetic animals were assumed to produce only one call each. Moreover, the maximum detection distance of the calls was set to be r_d = 5 km. Again, in a realistic and complex ocean environment detection distances are not known with certainty from data recorded at single instruments. A single frequency call was assumed at 20 kHz and a very simple detection procedure was adopted. All animals within 5 km of the hydrophone are detected with probability 1. The remaining animals outside r_d were not detected and hence the probability of detection was set to 0.

Having created a synthetic data set, the next step was to apply Eq. 1 to estimate density. Here, the assumptions made above for the creation of the data set were considered to be unknown. However, other assumptions needed to be made for the estimation problem. First, it was assumed that all animals were detected with certainty within some distance from the hydrophone and hence, the rate of false positive detections (c in Eq. 1) equaled 0. Assuming the hydrophone data were run through a generic detector, the number of detections was taken to be n_c = 54. It is noted that this figure corresponds to the number of animals within 5 km from the hydrophone, but this distance is not known for simulation purposes and some value is assumed. The maximum detection distance where all animals were sure to be detected was hence assumed to be w = 5 km. Parameters T and r were taken to be constants equal to 1 and hence, the average probability of detection, $\hat{P}$, was the only variable left that needed to be estimated.

The average probability of detection was estimated by employing a Monte Carlo simulation. The method was preferred instead of using available analytical formulas for the probability of detection simply to show how the problem would be solve in a more complicated scenario, discussed in Sec. IV. For the Monte Carlo run, 5000 animals were uniformly distributed inside the assumed circular area of radius w = 5 km. Transmission loss was calculated from each sampled animal position to the hydrophone using Eqs. 2 and 3. Two frequencies were used in TL calculations, 20 and 40 kHz. Calculations at 20 kHz were performed as a sanity check. Detections were considered as follows. If the calculated TL was bigger than TL at 5 km, then the sampled animal was not detected and the probability of detection was set to 0. Otherwise, a detection was considered and the probability of detection was set to 1. The estimated average probability of detection was computed by taking the average of all probabilities of detections of the 5000 samples.

B. Results

The number of animals from the original data set that were within 5 km from the hydrophone location, and hence considered detected, was n_c = 54. The true probability of detection P is defined for this case as the ratio between the true detection area ($\pi\ 5^2$ km^2) and the assumed detection area ($\pi\ 5^2$ km^2), or P = 1. The true density is given by the ratio between the total number of animals by the total circular area of ocean. The total number of animals corresponds to the number of points originally distributed and thus equals 1000. The total area is given by $\pi\ R^2$ (km^2) and therefore, D = 1000/$\pi\ 20^2$ = 0.796 animals/km^2.

When the frequency used to create the synthetic data was used in the Monte Carlo simulation, the estimated average probability of detection $\hat{P} = 1$, as expected. The estimated density from Eq. 1 becomes $\widehat{D} = 54/\pi\ 5^2 = 0.6875$ animals/km^2, which is close to the true density. The observed difference between the estimated and true densities is due to the random distribution of points used to create the synthetic data. If another data set were to be created, another n_c value would have been found, albeit close to the current value. Consequently, the density estimate would also have been different.

On the other hand, by using a different frequency in the Monte Carlo simulation (i.e. 40 kHz) yielded results that did not agree with the true values. For f = 40 kHz, $\hat{P} = 0.1866$, and $\widehat{D} = 3.6846$. The range of detection r_d for this frequency (found using Eq. 2 and 3) drops to 2.16 km and the bias in the probability of detection can be computed analytically. The bias in the estimation of P is given as the difference between its expected and true values, or $B(\hat{P}) = E[\hat{P}] - P = (\pi(2.16^2)/\pi 5^2) = 0.1866 - 1 = -0.8134$, or – 81%. The negative bias indicates that the average probability of detection was underestimated and consequently the density estimate was overestimated. This result was expected since the detection radius is smaller for higher frequencies and therefore, the number of animals detected is also reduced.

Another, perhaps more intuitive way to think about the probability of detection P is as a ratio between the true detection area and the assumed detection area ($P = \pi r_d^2/\pi w^2$). For this simple exercise that would give $\widehat{D} = n_c/\pi r_d^2$, or the number of animals per area, which is defined by the maximum range of detection. As mentioned previously, one of the complicating factors in the single sensor analysis is not knowing what the true detection range r_d really is. Therefore, one must use computational methods to estimate it as a probability of detection.

The purpose of this simple example was to show the impact of frequency on density estimates through well-known high-frequency propagation properties. By considering only the center frequency of a broadband signal to estimate average probability of detection can yield erroneous density estimates, which is dependent on the size of the detection circle, and consequently the number of animals inside it. Hence, *TL* modeling of higher frequencies will underestimate the probability of detection and overestimate population density estimates. Next, a more complex scenario is examined with a synthetic data set containing highly broadband signals.

IV. DENSITY ESTIMATION USING BROADBAND PROPAGATION

For this next computational experiment, a more realistic synthetic data set was created by convolving calculated arrivals with a synthetic source function resembling an echolocation click in terms of bandwidth. The single-sensor density estimation methodology was then applied to the synthetic data as it would be done for a real data set. A different approach for simulating broadband SNR of received clicks in the Monte Carlo run is also suggested. The procedure and its results are detailed below.

A. Construction of Synthetic Data Set

A circular area of radius 8 km was assumed inside which 100 synthetic *whales* were uniformly distributed as before (Sec. III). Next, the Bellhop ray tracing model (Porter and Bucker, 1987) was used to calculate ray arrival times and amplitudes from each of the

100 whales to the center of the area, i.e., the assumed sensor location. An isospeed waveguide with sound speed of 1500 m/s and a flat bottom at 2000 m were assumed in the simulation. Bottom parameters used as input to Bellhop were sound speed of 1600 m/s, bottom density of 1.5 kg/m^3, and bottom attenuation of 0.2 dB/m. Ray arrival information was calculated by assuming the whales were at the same depth of 600 m, which is consistent with depths where many cetacean species produce echolocation clicks. The receiver was placed at 620 m, similar to the HARP deployment depth (Sec. II).

Next, a broadband source function was approximated by a 500-point Tukey window with sampling frequency of 200 kHz and a peak-to-peak source level of 205 dB re 1 μPa. The resulting transient signal was further bandpass filtered between 10 and 60 kHz by applying a third-order Butterworth filter. The final waveform and its spectral content are shown in Fig. 4. This approximates a broadband echolocation click.

Frequency-dependent attenuation in the form $\exp(-\alpha r)$ was applied to the spectrum of the signal at each animal location. In the attenuation term, α has units of nepers/m and is given by Eq. 3 divided by $20\log(e)$ x 1000 (Jensen *et al.*, 2011), and r is distance from the recording sensor in meters. Attenuated signal spectra were then inverse Fourier transformed back to the time domain and convolved (Siderius and Porter, 2008) with the respective ray arrivals for each location. Noise was also added to the synthetic signals in the time domain to obtain corresponding received signals. Ambient noise data were extracted from spectrograms of the Hawai'i data set (see Sec. II) at several periods when no other sounds, such as ship noise or biological sounds, were present. The average of all noise samples per frequency bin was computed. Finally, the noise spectrum was scaled to be in agreement with typical noise levels from the Wenz curves (Wenz, 1962) for a sea state 3 (Fig. 5).

The received signal from each whale location corresponded to a 14-second long data sequence containing noise and ray arrivals convolved with a source function. All received signals were combined by sequentially placing each of the 100 14-second data segments into a single sound file for processing. Here, two synthetic data files were con- structed, one containing only first arrivals (without multipath) and the other considering all arrivals (with multipath). Details of the data analysis are presented next.

B. Synthetic Data Analysis

Data analysis for the cases with and without multipath followed the same procedure as if performed on real measured data. The first step was to manually annotate all visually detectable synthetic clicks on spectrograms of the data. These were plotted with the aid of the MATLAB-based application *Osprey* (Mellinger and Clark, 2006) by using a 512-point Hamming-windowed FFT with 50% overlap applied to every 0.5 s of data. Figure 6 shows both waveform and spectrogram of 2 s of synthetic data containing first, second, and third (very faint) arrivals from a synthetic whale at very close range (< 1 km). With the aid of *Osprey*, a box was manually placed around each click, from slightly before its onset to slightly after its end time. All start and end times were then saved to a log file. Manual detection yielded a total of 55 synthetic clicks if only first arrivals were considered and 136 clicks if multipath arrivals were also included. At closer ranges, of less than 1.5 km, up to six different arrivals could be manually detected from the synthetic data set. Detectable arrivals decreased in number to at most two at longer ranges.

Next *Osprey* was further used to measure the power of each click and of the average background noise. The mid-point of each manually detected click was found using the manually annotated start and end times. Half the synthetic click duration (total duration t = 0.0025 s) was subtracted and added to the start and end times, respectively. This step was necessary to ensure that power was measured in a reliable time window (enough time bins) containing the click. Three distinct frequency bands were considered for the measurements. The first was a 5 kHz frequency band centered on the synthetic click's center frequency of 35 kHz. The second included the click's full bandwidth from 10 to 60 kHz. The third frequency band was also a 5 kHz narrowband but centered at 20 kHz. Power measurements of background noise were performed in the same three frequency bands for a few time windows of distinct lengths in order to obtain a mean noise power. The mean noise power value was then used along with the measured click power to compute the SNR of each manually detected click. SNRs from this step are considered as *ground truth* and provide the basis to build the detection function.

The synthetic data was then run through a simple energy sum auto-detector with the aid of the software *Ishmael* (Mellinger, 2001). The same frequency bands used to calculate the SNR of manually picked clicks were also used in the auto-detections. Spectrogram parameters in *Ishmael* were set to be the same as those used in *Osprey* to make manual detections. The detection threshold for each frequency band was chosen such that there were no false positive detections, hence simplifying density calculations by making $c = 0$ in Eq. 1. The total number of auto-detected clicks in the three different frequency bands is presented in Table I.

In order to construct the detection function, i.e., the curve of probability of detection as a function of SNR, a binary data set was created from the comparison between auto and manual detections. If a manual detection was also picked up by the auto-detector, then a value of 1 was assigned to it, otherwise it was assigned a value of 0. This binary data along with the SNR information from manual detections was ran through a data regression routine written in the R language (The R Foundation for Statistical Computing, 2015). The R routine created 10 SNR value bins between the minimum and maximum SNRs measured from the data. A binomial probability was computed based on the number of SNR values, or detections, in each bin. A generalized additive model (GAM) (Wood, 2006) was then used to fit a curve to the estimated probability of detection. Finally, a logit (inverse logistic) function was used to link the GAM predictions to 300 SNR samples uniformly chosen between the minimum and maximum measured values. Five thousand random realizations of the GAM fit to the data were created and randomly sampled in the Monte Carlo simulation to estimate the average probability of detection $\hat{P}$ of Eq. 1 (Fig. 7).

C. Estimating the Average Probability of Detection

The passive form of the sonar equation (Urick, 1983) given by,

$$\text{SNR} = SL - TL - NL, \tag{4}$$

was used to calculate the SNR of five thousand click realizations. In Eq. 4 *SL* corresponds to source level (defined as the sound pressure level in dB re 1 μPa at a range of 1 m), *TL* to transmission loss (dB re 1 μPa), and *NL* to noise level (spectrum level in dB re 1 μPa normalized to 1 Hz band). Each realization, corresponding to a 2D whale location, was randomly sampled inside a circular area of radius 8 km. Before properly addressing the broadband nature of clicks, the following analysis considered the synthetic click's center frequency, as had been done previously (Küsel *et al.*, 2011).

Transmission loss was obtained from the incoherent solution of Bellhop calculations. Source frequency was set at the center frequency (35 kHz) for the cases when data analysis was performed at the 32.5–37.5 kHz and 10–60 kHz bands. When data analysis was performed at the 17.5–22.5 kHz band, *TL* was calculated at 20 kHz, or the center of the respective band. Input parameters to the Bellhop ray model were the same as be- fore (Sec. IV.A) and *TL* values were extracted at a single receiver depth of 600 m at the corresponding distances of each of the five thousand realizations. For consistency's sake, source level (*SL*) and noise level (*NL*) were taken as spectral levels at the same frequency as *TL* calculations, i.e., at 35 and 20 kHz. The value of source level at both frequencies was measured to be 134 dB re 1 μPa^2/Hz (Flat spectrum, see Fig. 4). The noise level at the click's center frequency of 35 kHz was taken to be approximately 36 dB re 1 μPa^2/Hz, and 43 dB re 1 μPa^2/Hz at 20 kHz (Fig. 5).

Even though a synthetic click is considered here, implying no uncertainties in *SL*, the uncertainty in the detection function is still accounted for in an outer loop of the Monte Carlo run. In other words, 300 realizations of the detection function were sampled, and for each of those 5000 realizations of SNR were estimated. Calculated SNRs were compared against a detection curve and a value of probability of detection was extracted. However, if a simulated SNR was lower than the lower SNR measured from the synthetic data, detection probability was taken as zero. On the other hand, if the simulated SNR was higher than the maximum observed in the synthetic data, detection probability of the maximum observed SNR was taken. The average probability of detection was obtained by calculating the mean of probabilities of detection of all 300 by 5000 realizations. Results of $\hat{P}$ as well as some of its statistics are shown in Table I for all cases considered. The standard error (SE) of $\hat{P}$ is given by the square root of its variance. The coefficient of variation (CV) is the SE of $\hat{P}$ divided by $\hat{P}$ and is given as a percentage.

It is worth noting that for the case of the narrowband centered on the click's center frequency and disregarding multipath arrivals, the data regression yielded an exact step function. It is also observed that regardless of the inclusion of multipath arrivals, the bandwidth used in the data analysis did not influence $\hat{P}$ substantially. Looking at Fig. 7, plots (a) and (c), it is noted that despite the different bandwidth used in the auto-detection process and the qualitative differences between detection functions, similar estimates were obtained for $\hat{P}$, 0.1216 and 0.123. Furthermore, the inclusion of multipath arrivals in the analysis only contributes with more samples with which to construct the detection function (Fig. 7(c)). In other words, small fluctuations in the detection function did not cause big variations in the average probability of detection. However, the frequency used in TL calculations seemed to have the biggest impact on $\hat{P}$, as could be observed from the results presented in Table I for the 17.5□22.5 kHz frequency band, in which TL was computed at 20 kHz.

D. Density Estimation Results

As 100 animals were simulated across a circular area with radius 8 km, the true density was equal to 497 animals per 1000 km^2. Density was estimated from the synthetic data set with emphasis on the effect of detections (n_c) and their average probability ($\hat{P}$) on the results, therefore all other parameters in Eq. 1 were considered as fixed and non-varying. There were no false positives, so $c = 0$. The maximum detection range, w, was 8 km and the cue production rate, r, corresponded to 1 since it was assumed that each point only produced a click once. The time period T was also taken to be equal to 1. Density estimates along with coefficients of variation and 95% confidence intervals are shown in Table I for the three frequency bands considered, with and without multipath. Coefficients of variation of density estimates were calculated by taking the square root of the sum of the squared coefficients of variation of $\widehat{P}$ and n_c.

Using a synthetic data set, albeit simplified, it was possible to observe the magnitude of the error introduced to density estimates by using the sonar equation to estimate received SNRs and hence their probabilities of detection. The density estimates produced by assuming the full click bandwidth in the auto-detections, but by using the center frequency to calculate TL used in the passive sonar equation (Eq. 4), were 2.65 (no multipath) and 6.5 (with multipath) times higher than the true density.

On the other hand, considering the two narrowband (5 kHz bandwidth) cases, it was observed that the density estimated for the one centered around 20 kHz and disregarding multipath arrivals ($\widehat{D}$ = 440) gave the closest result to the true density (D = 497). Since the chosen synthetic signal was truly broadband in nature, with a flat spectrum from 10 to 60 kHz, using any narrow band along the broader bandwidth should yield somewhat similar results, closer to the true value. When considering multipath arrivals in the analysis, it was observed that auto-detections were roughly two fold compared to the no multipath cases. If first arrivals were detected, most likely the second was also detected. In the few cases where the whale was closer to the sensor location, up to four multipath arrivals were auto-detected. This explains the density estimate results in the multipath case being roughly double of the expected value.

By reducing the broadband problem to that of a narrow band one in the detection process also reduces the complication of simulating broadband signals using the CW passive sonar equation. Even though the center frequency carries most of the energy in the click, it will also be attenuated more rapidly with distance and for a simple synthetic data set there was barely enough detections to construct a detection function. That was the reason for picking a 5 kHz bandwidth centered at a lower frequency (20 kHz) within the click's broad bandwidth. However, this may not pose a problem when dealing with real data sets where thousands of detections are usually made. It was also observed that multipath detections can cause an increase in the estimates. For real data sets it may be very hard or impossible to distinguish between multipath arrivals. Given the issues related to using the sonar equation, especially for a very broadband source, a different approach explained next, is tested.

E. Calculating SNR from Ray Arrival Information

Equation 4 is customarily used in the detection analysis of single frequency signals. The biggest issue in using this equation for single sensor density estimation studies is the fact that marine mammal calls can be very broadband in nature. When using the passive sonar equation to estimate SNR, bandwidth is not the only issue to be considered. The TL term, which corresponds to transmission loss, in fact represents the sum of all contributing multipath arrivals. Hence, when calculating the SNR of each click realization, one is actually calculating the SNR of all multipaths summed together. However, each click in the data set corresponds to one ray arrival, just as was realized when creating the broadband synthetic data set.

An alternative to using the passive sonar equation mirrors the procedure used to create the synthetic data set. Ray arrival amplitude and delay information is calculated for each click realization distributed inside a circular area. Frequency-dependent attenuation is added to the arrivals which are then convolved with a source function, i.e. a strong on-axis click representative of the species of interest. Background noise is added to the convolved signals in the time domain, forming a synthetic received signal. Each synthetic received signal is then stacked together to create a sound file. In this manner, SNR can be estimated in a similar fashion as is done when measuring SNR from a data set, i.e., from the data spectrogram. In this step, care must be taken so that power measurements agree with data power measurements performed earlier in the analysis.

The above approach was also applied to estimate the average probability of detection from the broadband synthetic data set. As a sanity check, the exact 100 animals of the synthetic data set were used in the Monte Carlo simulation. Considering sampling variations, a density estimate very close to the true density would be expected. This exercise was performed by considering only first arrivals and by using the full click bandwidth (10 to 60 kHz) to estimate SNR and hence, $\hat{P}$. The sanity check yielded 529 animals per 1000 km^2, which is 1.06 times the true density of 497 whales per 1000 km^2. It was observed that a small amount of error was introduced from differences in time windows used to measure the SNR of clicks. This in turn yielded small variations in SNR as originally measured by *Osprey* which, when compared against the almost step-like detection function, probably caused some variations in probability of detection.

Next, considering 5000 click realizations inside the circular area of radius 8 km, the Bellhop ray tracing model was used in a single run to calculate arrival amplitudes and time delays, as before. The calculation of arrivals was performed with the same parameters used for the creation of the synthetic data set. Following the same steps described in Subsection IV.A, frequency attenuation was added to the calculated arrivals and the results convolved with the synthetic source function. Finally, noise was added to the attenuated signals. The SNR of each simulated received signal was measured from its spectrogram, where the onset of each click was obtained from the time delay given by the ray tracer, and the pulse duration was added to it. The spectrogram was calculated with the same parameters used in the data analysis, i.e., a 512-point FFT, with 50% overlap and using a Hamming window. Considering only the first arrivals, their power was calculated by summing the frequency contributions from 10 to 60 kHz. A two-second window chosen sometime after the first arrival was used in the same way to measure the power of background noise. Signal-to-noise ratio was then calculated as the difference in dB between click and noise powers. The remainder of the analysis was the same as that described when using the sonar equation, but the detection function used in this case was the one created using the entire click bandwidth.

Results of the Monte Carlo run for 300 detection functions by 5000 whale locations, yielded an average probability of detection, $\hat{P}$, equal to 0.3012 (SE = 0.0057; CV = 1.9%). Using this figure in Eq. 1 with 32 auto-detected clicks (Table I), yielded a density estimate of 528 whales per 1000 km^2 (CV = 18%; 95% CI = 421-664). Such result is very encouraging and a slightly better estimation than that given by the narrowband approximation (5 kHz bandwidth centered at 20 kHz). However, it is worth noting that multipath arrivals are not being considered in this comparison.

As mentioned earlier, extra arrivals resulting from boundary reflections simply add to the pool of SNRs used to construct the detection function as well as to the number of detected clicks. The first has little implication to the final density estimation whereas the latter is an important factor. The number of multipath arrivals however, is not uniformly distributed and will decrease in number the further the animal is from the receiver. Therefore, including a set number of arrivals in the Monte Carlo run will bias the final average probability of detection. In fact, it would represent an increase in the number of realizations and a decrease in $\hat{P}$, since there would be more samples representing zero probability of detection to be averaged out. Ideally, the average probability of detection would be estimated considering only first arrivals, and the density estimate would be multiplied by a scaling factor to take multipath detections into account. Using the value of $\hat{P}$ = 0.3012 in Eq. 1 but taking n_c = 80 auto-detections, which includes multipath arrivals, yields a density estimate of $\hat{D}$ = 1321 animals per 1000 km^2. This corresponds to 2.5 times the density obtained from first arrivals only. When dealing with real data sets, a similar exercise involving the simulation of synthetic data that is run through the same detector as the real data could yield this scaling factor. Alternatively, a scaling factor could be derived from the Monte Carlo run by incorporating the detector's threshold and analyzing how many clicks would be detected if only first arrivals were considered or if multipath was also considered.

Another advantage of working with arrivals, instead of using the sonar equation, is that there is no need to know about source level distributions. However, a source function representative of the click is needed to convolve with arrival information. An on-axis source function could be potentially extracted from the data set being analyzed. Also, instead of using a model for the animal's beampattern (e.g. circular piston model), if such information is available, the vertical beampattern can be readily incorporated into the ray tracer model for the calculation of arrivals. Azimuthal beampattern can also be easily taken into account.

V. DISCUSSION AND CONCLUSIONS

In this work the population density estimation methodology for data sets recorded on single, fixed instruments was revisited, keeping in mind the broadband nature of many cetacean echolocation clicks. These types of marine mammal calls are suitable and usually preferred for density estimation studies. For single-sensor analysis, the most important parameter of the density estimator equation is the average probability of detection, $\hat{P}$. Its estimation can be quite different if the sound source is narrow or broadband in nature. For broadband sources, it was shown that the analysis should take into consideration the frequency content of the sound in order to obtain a reliable estimate of $\hat{P}$, and hence density.

The error in estimating the probability of detection by simulating a source's frequency band by using its center frequency was illustrated through a simple example using a synthetic data set (Sec. III). Using the center frequency changes the size of the detection region and consequently the number of animals inside this region that could be detected. Considering a detection region that is smaller than the true one will produce a lower average probability of detection. Consequently, density estimates will be higher than expected.

Although measurements are often preferred, it was shown in Sec. IV that simple designed simulations can provide useful information regarding the estimation of $\hat{P}$. In fact, by making the correct assumptions in the calculation of $\hat{P}$ yields density estimates that are very close to the true values. Furthermore, true values are known *a priori*, which gives a good way to check results of simulations. Hence, an alternative approach was proposed and tested for estimating the SNR of broadband marine mammal calls. Instead of using the sonar equation, which is an inherent technique for single frequency sounds, SNR was estimated from arrival information convolved with the source function representative of the call of interest to obtain a simulated received signal. Such approach reflects more closely what is observed in real data sets, i.e., each sound represents a single arrival. Moreover, transmission loss, which is used in the sonar equation, is in fact the sum of all arrivals. However, one detected click (broadband signal) corresponds to a single arrival.

Acknowledgments

This research was supported under the Office of Naval Research, Marine Mammal and

Biology Program. The authors would like to thank Erin Oleson, from the Pacific Islands Fisheries Science Center, and Simone Baumann-Pickering and Anne Simonis, from Scripps Whale Acoustic Lab, for providing the data set used in this study as well as invaluable information on detections of false killer whales. We thank Laura Koeppler for sharing a click waveform from the captive false killer whale Kina. We thank Evan Howell, from the Pacific Islands Fisheries Science Center, Ecosystems and Oceanography Division, for kindly sharing oceanographic data collected off the Kona coast. We are also grateful to Robin Baird for invaluable information and references on false killer whale populations in Hawai'i. This is PMEL contribution number 4524.

References

Ainslie, M.A. (**2010**). *Principles of sonar performance modelling* (Springer-Verlag Berlin Heidelberg), pp. 53–122.

Ainslie, M.A. (**2013**). "Neglect of bandwidth of odontocetes echo location clicks biases propagation loss and single hydrophone population estimates," J. Acoust. Soc. Am. **134**, 3506–3512.

Jensen, F.B., Kuperman, W.A., Porter, M.B., and Schmidt, H. (**2011**). *Computational ocean acoustics* (Springer), pp. 35-36.

Küsel, E.T., Mellinger, D.K., Thomas, L., Marques, T.A., Moretti, D., and Ward, J. (**2011**). "Cetacean population density estimation from single fixed sensors using passive acoustics," J. Acoust. Soc. Am. **129**, 3619–3622.

Marques, T.A., Munger, L., Thomas, L., Wiggins, S., and Hildebrand, J.A. (**2011**). "Estimating North Pacific right whale *Eubalaena japonica* density using passive acoustic cue counting," Endang. Species Res. **13**, 163–172.

Marques, T.A., Thomas, L., Ward, J., DiMarzio, N., and Tyack, P.L. (**2009**). "Estimating cetacean population density using fixed passive acoustic sensors: An example with Blainville's beaked whales," J. Acoust. Soc. Am. **125**, 1982–1994.

Marques, T.A., Thomas, L., Martin, S.W., Mellinger, D.K., Ward, J.A., Moretti, D.J., Harris, D., and Tyack, P.L. (**2013**). "Estimating animal population density using passive acoustics," Biol. Rev. **88**, 287–309.

Mellinger, D.K. (**2001**). "Ishmael 1.0 User's Guide," NOAA Technical Memorandum OAR PMEL-120, NOAA/PMEL/OERD, 2115 SE OSU Drive, Newport, OR 97365-5258.

Mellinger, D.K. and Clark, C.W. (**2006**). "MobySound: A reference archive for studying automatic recognition of marine mammal sounds," Appl. Acoust. **67**, 1226–1242.

Porter, M.B. and Bucker, H.P. (**1987**). "Gaussian beam tracing for computing ocean acoustic fields," J. Acoust. Soc. Am. **82**, 1349–1359.

R Foundation for Statistical Computing (**2009**). "R: A Language and Environment for Statistical Computing", URL `http://www.R-project.org`.

Roch, M.A., Klinck, H., Baumann-Pickering, S., Mellinger, D.K., Qui, S., Soldevilla, M.S., and Hildebrand, J.A. (**2011**). "Classification of echolocation clicks from odontocetes in the Southern California Bight," J. Acoust. Soc. Am. **129**, 467–475.

Siderius, M. and Porter, M.B. (**2008**). "Modeling broadband ocean acoustic transmissions with time-varying sea surfaces," J. Acoust. Soc. Am. **124**, 137–150.

Soldevilla, M.S., Henderson, E.E., Campbell, G.S., Wiggins, S.M., Hildebrand, J.A., and Roch, M.A. (**2008**). "Classification of Risso's and Pacific white-sided dolphins using spectral properties of echolocation clicks," J. Acoust. Soc. Am. **124**, 609–624.

Thomas, L. and Marques, T.A. (**2012**). "Passive Acoustic Monitoring for Estimating Animal Density," Acoustics Today **8**, 35–44.

Urick, R. J. (**1983**). *Principles of Underwater Sound* (McGraw-Hill Ryerson, Limited). pp. 17–30, 111.

von Benda-Beckmann, A.M., Lam, F.P.A., Moretti, D.J., Fulkerson, K., Ainslie, M.A., van IJsselmuide, S.P., Theriault, J., and Beerens, S.P. (**2010**). "Detection of Blainville's beaked whales with towed arrays," Appl. Acoust. **71**, 1027–1035.

Wenz, G.M. (**1962**). "Acoustic Ambient Noise in the Ocean: Spectra and Sources," J. Acoust. Soc. Am. **34**, 1936–1956.

Wiggins, S.M. and Hildebrand, J.A. (**2007**). "High-frequency Acoustic Recording Package (HARP) for broad-band, long-term marine mammal monitoring," in *International Sym-*

posium on Underwater Technology 2007 and International Workshop on Scientific Use of Submarine Cables and Related Technologies 2007. IEEE. pp. 551–557.

Wood, S.N. (**2006**). *Generalized Additive Models: An Introduction with R.* (Chapman and Hall/CRC Press, Boca Raton, FL), p. 262.

TABLE I. Results from auto detection and density estimation, including statistics, for the broadband synthetic data with and without multipath, and three frequency bands, two narrow and one broad band. For the first two frequency bands, *TL* was calculated at 35 kHz, and for the third band *TL* was calculated at 20 kHz. Density estimates are given in number of whales per 1000 km^2. (SE = standard error; CV = coefficient of variation; CI = confidence interval)

Freq. Band	Auto Dets.	$\hat{P}$	SE ($\hat{P}$)	CV ($\hat{P}$)	$\hat{D}$	CV ($\hat{D}$)	95% CI ($\hat{D}$)
	Without Multipath						
$32.5 \sim 37.5$ kHz	8	0.1173	7.42×10^{-6}	0.0063%	339.2	35.35%	218-528
$10 \sim 60$ kHz	32	0.1210	6.89×10^{-4}	0.57%	1316	17.7%	1049-1651
$17.5 \sim 22.5$ kHz	27	0.3052	5.88×10^{-4}	0.19%	440	19.25%	344-563
	With Multipath						
$32.5 \sim 37.5$ kHz	23	0.1216	5.13×10^{-4}	0.42%	941.1	20.86%	721-1228
$10 \sim 60$ kHz	80	0.1230	8.84×10^{-4}	0.72%	3235	11.2%	2801-3737
$17.5 \sim 22.5$ kHz	58	0.3021	0.0029	0.96%	955	13.17%	806-1131

List of Figures

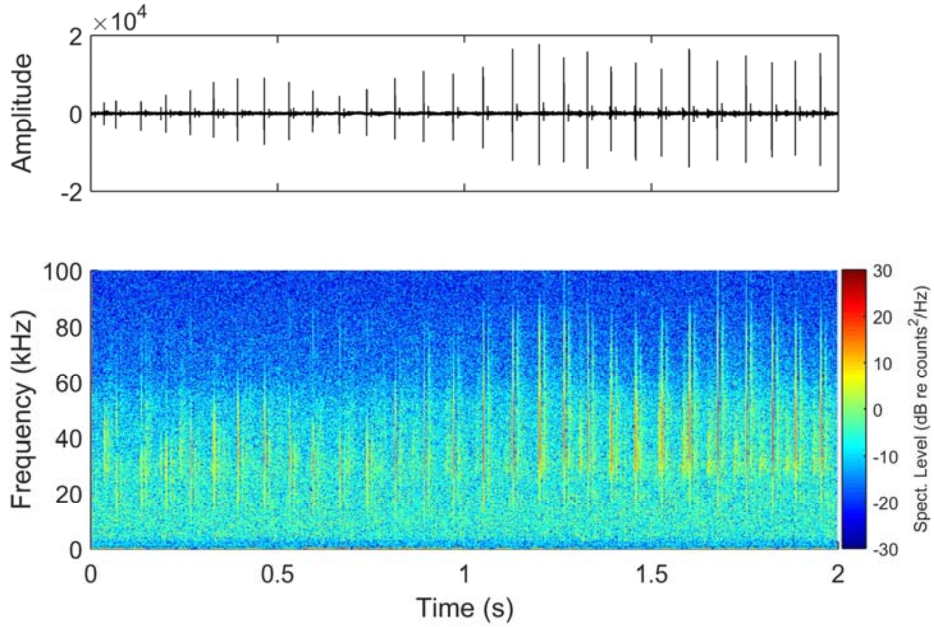

×10^4
Amplitude
Frequency (kHz)
Spect. Level (dB re counts^2/Hz)
Time (s)

FIG. 1.

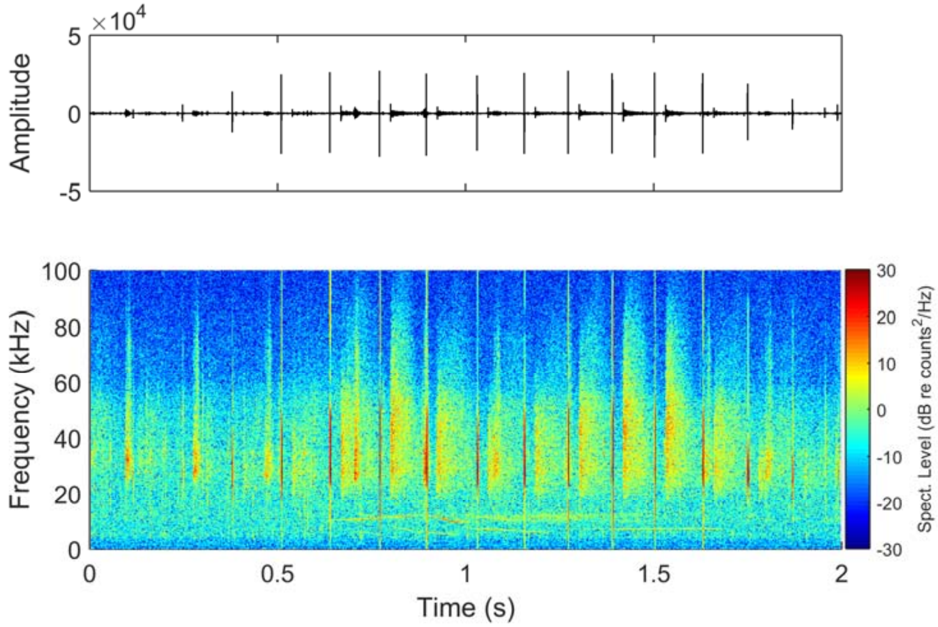

×10^4
Amplitude
Frequency (kHz)
Spect. Level (dB re counts^2/Hz)
Time (s)

FIG. 2.

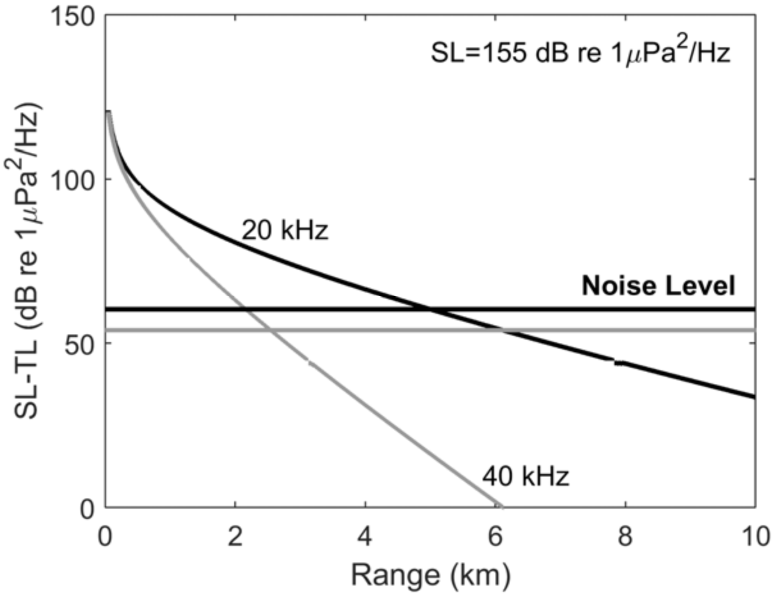

150
100
50
0
SL=155 dB re 1μPa^2/Hz
SL-TL (dB re 1μPa^2/Hz)
20 kHz
Noise Level
40 kHz
0
2
4
6
8
10
Range (km)

FIG. 3.

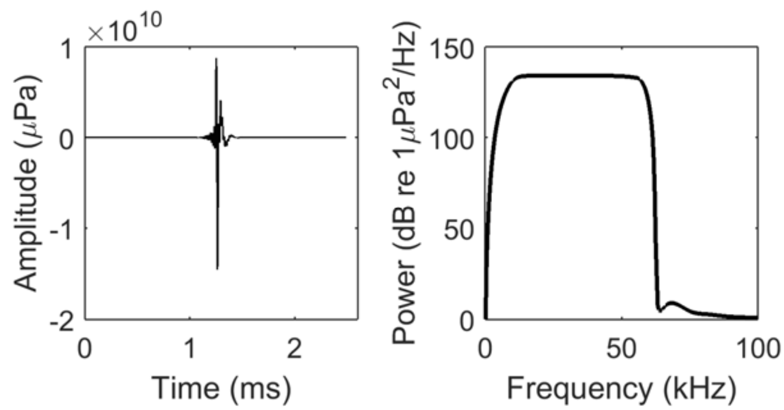
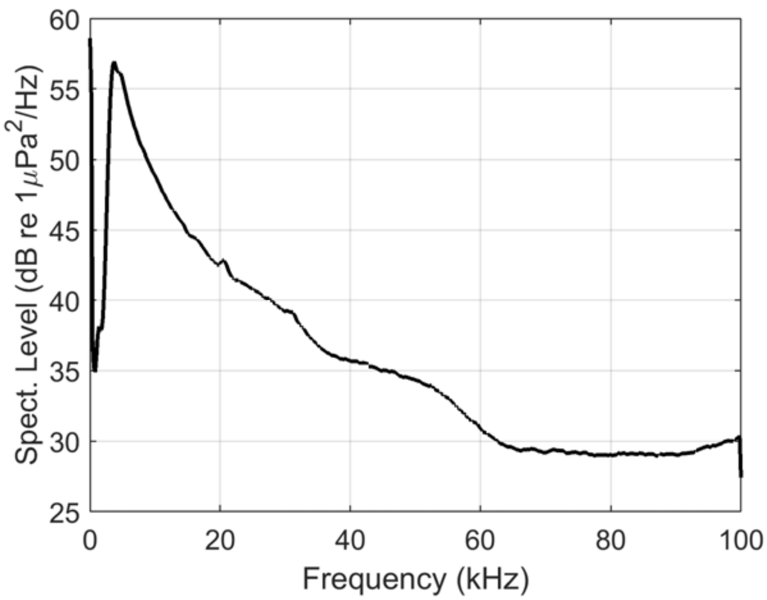

$\times 10^{10}$
1
0
-1
-2
Amplitude (μPa)
0
1
2
Time (ms)
150
100
50
0
Power (dB re 1μPa^2/Hz)
0
50
100
Frequency (kHz)

FIG. 4.



Spect. Level (dB re 1μPa²/Hz)
60
55
50
45
40
35
30
25
0
20
40
60
80
100
Frequency (kHz)

FIG. 5.

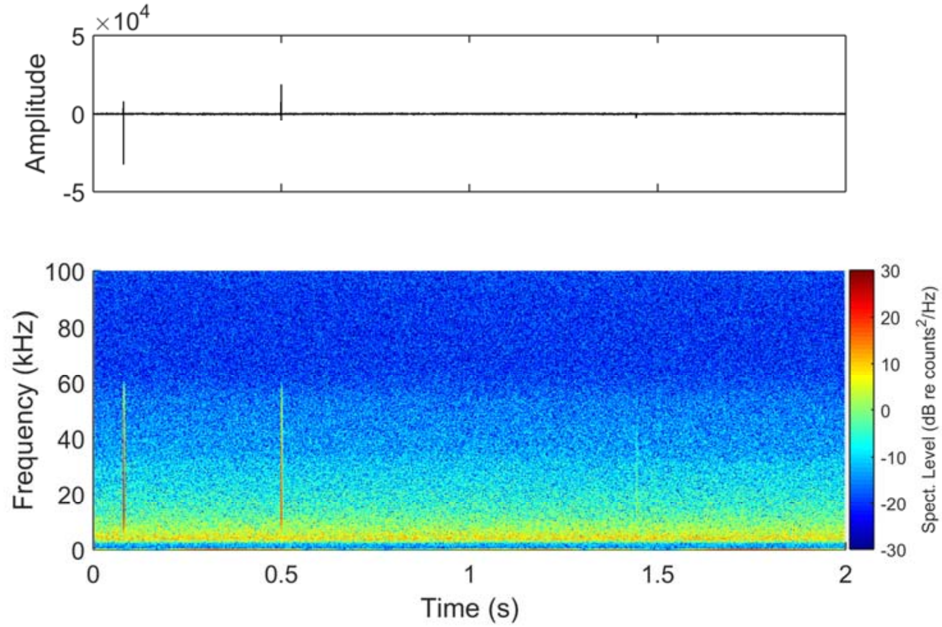

×10^4
5
0
-5
Amplitude
100
80
60
40
20
0
Frequency (kHz)
0
0.5
1
1.5
2
Time (s)
30
20
10
0
-10
-20
-30
Spect. Level (dB re counts^2/Hz)

FIG. 6.

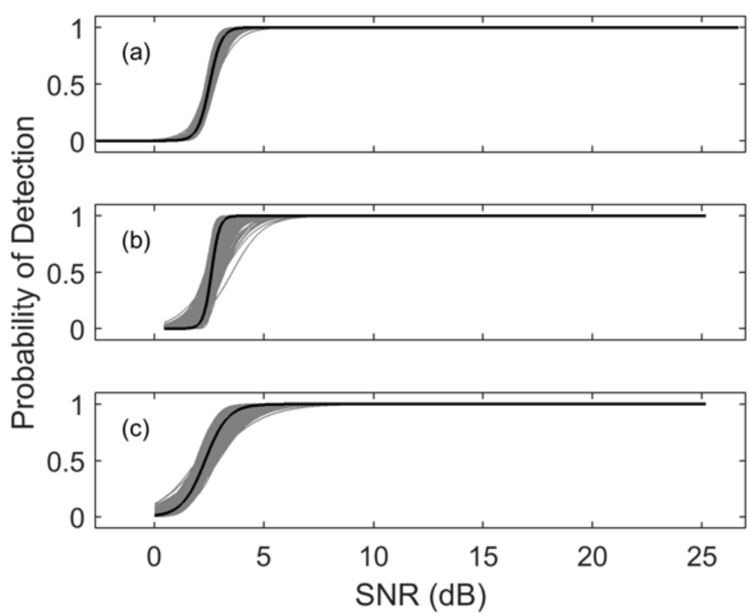

(a)
(b)
(c)
1
0.5
0
Probability of Detection
0
5
10
15
20
25
SNR (dB)

FIG. 7.