Estimating minke whale (Balaenoptera acutorostrata) boing sound density using passive acoustic sensors

Article *in* Marine Mammal Science · January 2013

DOI: 10.1111/j.1748-7692.2011.00561.x

Estimating minke whale (*Balaenoptera acutorostrata*) boing sound density using passive acoustic sensors

Stephen W. Martin
Space and Naval Warfare Systems Center Pacific
53560 Hull Street, San Diego, CA 92152. U.S.A.

Tiago A. Marques
Centre for Research into Ecological and Environmental Modelling
University of St Andrews
The Observatory, Buchanan Gardens, St Andrews KY16 9LZ, Scotland
and
Centro de Estatística e Aplicações da Universidade de Lisboa, Bloco C6, Piso 4,
Campo Grande, 1749-016 Lisboa, Portugal

Len Thomas
Centre for Research into Ecological and Environmental Modelling
University of St Andrews
The Observatory, Buchanan Gardens, St Andrews KY16 9LZ, Scotland

Ronald P. Morrissey, Susan Jarvis, Nancy DiMarzio and David Moretti
Naval Undersea Warfare Center
1176 Howell Street, Newport, RI 02841. U.S.A.

David K. Mellinger
Cooperative Institute for Marine Mammal Studies
Oregon State University
2030 SE Marine Science Dr., Corvallis, OR 97365. U.S.A.

ABSTRACT

Density estimation for marine mammal species is performed primarily using visual distance sampling or capture-recapture. Minke whales in Hawaiian waters are very difficult to sight; however, they produce a distinctive “boing” call, making them ideal candidates for passive acoustic density estimation. We used an array of 14 bottom-mounted hydrophones, distributed over a 60 by 30 km area off Kauai, Hawaii, to estimate density during twelve days of recordings in early 2006. We converted the number of acoustic cues (*i.e.*, boings) detected using signal processing software into a cue density by accounting for the false positive rate and probability of detection. The former was estimated by manual validation, the latter by applying spatially explicit capture-recapture (SECR) methods to a subset of data where we had determined which hydrophones detected each call. Estimated boing density was 130 boings per hour per 10,000 km^2 (95% CI 104-163). Little is known about the population’s acoustic behavior, so conversion from boing to animal density is difficult. As a demonstration of the method, we used a tentative boing rate of 6.04 boings per hour, from a single animal tracked in 2009, to give an estimate of 21.5 boing-calling minke whales per 10,000 km^2.

Keywords: cue count, passive acoustic density estimation, passive acoustic monitoring, boing vocalization, spatially explicit capture recapture

INTRODUCTION

Minke whales (*Balaenoptera acutorostrata*) are one of the smallest baleen whales but one of the most abundant (Reeves *et al*. 2002). They are cosmopolitan, occurring in polar regions during the summer months and migrating to warmer tropical waters to breed in the winter (Stewart and Leatherwood 1985). As with many other cetacean species, density and abundance estimates, where available, are largely based on visual line transect surveys (*e.g.*, de Boer 2010). However, in some areas, the animals appear to be extremely cryptic to visual observation, making such methods infeasible. On the other hand, at least some component of the population produces a readily detectable call, making it potentially amenable to population estimation via passive acoustic methods. In this paper, we demonstrate the feasibility of such an approach, focusing on waters off Hawaii where there is a pre-existing hydrophone array. We utilize a novel method for converting the number of detected calls to a call density, and use a tentative estimate of call rate to show how this can be used to estimate animal density.

Based upon documented sightings, minke whales are suspected to occur in Hawaii only during winter and spring (Balcomb 1987, Caretta *et al.* 2005, Rankin and Barlow 2005, Rankin *et al.* 2007). Despite over 10,000 km of aerial line transects[1] for marine mammals around the island of Kauai, there have been no sightings of minke whales during these aerial surveys. However, if detectability were similar to that found for other visual surveys (*e.g*., figure 5 in Zerbini et al 2006) then the low sighting rate implies that density is very low. An

[1] J. Mobley personal communication, 2528 McCarthy Mall, Honolulu, HI, Feb. 2010

alternative explanation is that minke whales in Hawaiian waters are very cryptic to visual observation.

Boing sounds have been detected in waters near the Hawaiian Islands by the US Navy for decades (Wenz 1964). The boing sounds occur most often during February and March and have long been thought to be produced by whales (Thompson and Friedl 1982). The boing sound has also been described as having characteristics similar to the dwarf minke whale "star wars" sound (Gedamke *et al.* 2001). However, it was not until late 2002 that the boing sound was shown to be made by minke whales (Rankin and Barlow 2005). A recent minke whale sighting in Hawaiian waters on 27 April 2009 was facilitated by directing a surface vessel towards boing calls detected and localized using bottom mounted hydrophones at the US Navy Pacific Missile Range Facility (PMRF) north of Kauai, Hawaii[2] (see Methods for more details about the hydrophones and acoustic processing). During a 24-hour period starting at 0749 HST that day, over 15,000 automatic detections of boing sounds were obtained from the 14 hydrophones used at PMRF, and less than 8 percent were estimated to have come from the sighted individual (a single boing call from an animal may result in zero, one, or more automatic detections by a single hydrophone). The quantity of boings detected suggests that minke whales are more abundant in the area during late winter and early spring than previously believed. It also suggests that it may be possible to use the acoustic detections to estimate density, at least of the component of the population producing boings. The purpose of the boing call is currently unknown, although current

[2] Tom Norris personal communication, 517 Cornish Dr. Encinitas, CA 92024, April 2009.

thought is that it is related to breeding and may only be emitted by reproductively active males, much like Humpback song vocalizations[3]. Estimation of absolute density and abundance from passive acoustic detections recorded at static hydrophones is a rapidly developing field, and a number of approaches are now available (*e.g.*, Küsel *et al.* 2011, Marques *et al.* 2009, Marques *et al.* 2011*b*, Moretti *et al.* 2010). The method most suitable for a given situation depends on the target species' acoustic behavior, the available acoustic detection capabilities, and the auxiliary information available about the sound production and detection processes. In some rare instances such as a species with loud and frequent calls that occur in a dense array of hydrophones, it may be feasible to assume that all diving individuals within a known area can be detected and counted, and all those outside the area excluded (Ward *et al.* in press). Density estimation then simply involves dividing counts taken at a sample of temporal "snapshots" by the area surveyed and by an estimate of the average proportion of animals diving (obtained, for example, from focal follows or tagging studies). However, in most studies (including the one reported here), estimation is not so straightforward, for three reasons. Firstly, it is not possible to determine whether successive detected sounds come from the same animal; hence the input data is a count of detected "cues" (here boing sounds) on each hydrophone rather than detected animals. This can be accounted for by estimating the density of cues, and then dividing by an estimate of cue rate (derived again from focal follows, tagging studies, or other auxiliary information) to estimate animal density. Secondly, some cues detected by the automated

[3] Jay Barlow personal communication, 3333 N. Torrey Pines Ct., La Jolla, CA. September, 2009.

signal processing algorithms typically employed may be “false positives” – *i.e.*, they may not be sounds emitted by the target species, or they might be multipath arrivals of the same original sound. This can be accounted for if a sample of data can be accurately hand-validated, and an estimate of average false positive proportion derived. Thirdly, there will also be “false negatives” – *i.e.*, vocalizations produced within the study area that are missed. There are a variety of methods to estimate the proportion missed, also called the detection probability.

Here, we make use of the fact that minke whale boing calls are loud and sufficiently omnidirectional that individual calls can often be detected on multiple hydrophones, and that it is possible to associate detections from the same boing across hydrophones. The resulting information about which hydrophones detect each boing can be thought of as a “capture history”, analogous to that used in capture-recapture studies to estimate wildlife population size, with additional information about the location of each capture (*i.e.*, the location of each hydrophone that detected a boing). This opens the way to application of recently-developed spatially explicit capture-recapture (SECR) methods to estimate boing density. Such methods were originally developed for application in small mammal trapping studies (*e.g.*, Efford 2004), and have been applied to passive acoustic data on songbirds (Dawson and Efford 2009). Using a small, 60 minute, test data set of minke whale boing data from PMRF, Marques *et al.* (2011*a*) compared alternative statistical implementations of SECR. Here, we extend this work in four respects. Firstly, we use a much larger sample of 12 days of recordings (167 h of data) to derive detected boing counts (termed the detected boings data set). Secondly, we apply SECR to only a subset of these data to derive estimates of detection probability, and show how the information from this subset can be combined with the larger sample of count data to obtain a density estimate and corresponding variance. This is useful because processing the acoustic data to associate calls and creating the capture histories that are input to SECR is labor-intensive and hence becomes infeasible when recordings cover long time periods, as is often the case for static acoustic monitoring. Thirdly, our modeling of the SECR data accounts for potential variation in detection probability due to the varying depths of the PMRF hydrophones, and for the masking effect of the Hawaiian Islands. Fourthly, we assess the sensitivity of the density estimate relative to different human operators performing manual boing validations.

In the following we refer to “boing call density” for the density of minke whale boing calls (calls per unit area per unit time), which when divided by a “cue rate” (long term average calls per unit time per individual whale) leads to an estimate of boing calling minke whale density (whales per unit area). We use the above extensions to derive an estimate of the mean boing call density over the approximately 167 hours of recordings, with corresponding confidence limits. As with many studies, conversion of this to animal density is problematic because reliable estimates of mean cue rate are lacking. For minke whales the boing cue rate appears bimodal in nature with rates of both approximately 330 s and 30 s between calls (Thompson and Friedl 1982 and Rankin and Barlow 2005). A robust long term average cue rate therefore needs to properly represent the complexity of the cue rate. To illustrate the method, we use a tentative estimate of cue rate, based on a single animal tracked in 2009 emitting only the slower rate of calls, to derive a density estimate for minke whales producing the boing calls.

METHODS

Estimation of Density and Variance

Density estimation using acoustic cues has been applied to a beaked whale species (Marques *et al.* 2009). The basic premise is that the estimated density of animals is determined by a measure of the number of acoustic cues detected (n_c), corrected by the estimated average false positive proportion ($\hat{c}$), divided by the product of the study time (T), the estimated mean probability of detecting the cue over the area surveyed ($\hat{P}$), the area surveyed (A), and the estimated cue rate ($\hat{r}$).

Equation 1 $$\widehat{D}=\frac{n_c(1-\hat{c})}{T\times A\times\widehat{P}\times\widehat{r}}=\frac{n_c}{T}\frac{(1-\hat{c})}{\widehat{P}}\frac{1}{A}\frac{1}{\widehat{r}}=\frac{n_c}{T}\frac{(1-\hat{c})}{\hat{a}}\frac{1}{\widehat{r}}$$

Note that multiplying the area surveyed by the estimated mean probability of detection ($\hat{P}\times A$) leads to an estimated effective sampling area, $\hat{a}$, *i.e.,* the area such that on average one would detect as many sounds in the survey as if all sounds produced within $\hat{a}$ were detected (see Buckland *et al.* 2001 for an extensive discussion of interpretation of effective sampling area). While this is a relatively straightforward equation, obtaining estimates of the required random components can be difficult. The case study in Marques *et al.* (2009) relied on having digital time, acoustic, and depth tags, DTAGs (Johnson and Tyack 2003), on several beaked whales in order to estimate these parameters. The effort, time and cost involved in obtaining sufficient tag samples are high especially when one considers an elusive species such as minke whales in Hawaiian waters.

In this study we depart from the above paradigm. While one can frame the above estimator as being based on distance sampling cue-counting methods (Buckland et al, 2001), the one we use here is based on SECR methods, a rapidly evolving statistical technique (for details see, *e.g.*, Borchers and Efford, 2008; Royle 2009). SECR is well suited to obtaining the average probability of detection or effective sampling area using multiple acoustic hydrophone sensors, also termed proximity traps (Efford, Borchers and Byrom 2009). SECR methods are based on capture-recapture data, corresponding here to the detection of the same boing call on multiple hydrophones. The known location of the hydrophones gives rise to a spatial indexing of the boing detections as the sound propagates through the area. The method also requires underlying models for (1) how animals are detected (here how sounds are detected) as a function of distance from the detector (a “detection function” model) and (2) how animals distribute themselves in space. Regarding the detection function model, different detection functions were investigated: half normal (Hn), hazard rate (Hr), negative exponential (Ne) and cumulative lognormal (Ln). The use of hydrophone depth as a covariate in the scale parameter of the detection function, *i.e.*, allowing detectability to be a function of depth, was also attempted (using an identity link). The best fitting model for the detection function was selected using Akaike Information Criterion (AIC; Akaike 1985). We considered only models for which detection of a sound produced near the sea surface of the area directly over the hydrophone was certain ($g(0)=1$), which seems reasonable in the acoustic setting considered (see Discussion). Regarding the distribution in space, the likelihood used here assumes implicitly that the distribution of animals in space is a homogeneous Poisson process (*i.e.*, uniform density in space, with animal locations being independent of one another).

The analysis was implemented in the statistical software R (R Development Core Team, 2009) using the *secr* package (Efford 2009). This package implements computations on a discrete grid, called the habitat mask. The habitat mask must include all areas from where potential boings might have been detected. Here the habitat mask was constructed to cover the area out to 210 km from any hydrophone, which given minke boing source levels (*e.g.*, Thompson and Friedl 1982) and sound propagation characteristics in the area seems a very conservative value. The habitat mask also allows the user to easily define areas of non-habitat, where by assumption either no animals produce boings or boings produced in the masked area have zero probability of detection. Here we consider these to correspond to the island land areas, as well as the areas shielded from a given hydrophone by land (*cf*. Fig. 1).

Using SECR implies that the number of unique boings detected during the survey period must be quantified, which is not straightforward for two reasons:

(1) the existence of false positives, and (2) most boings are detected on multiple hydrophones, but without associations we do not know how many. Therefore, we not only need, as with previous studies (*e.g.*, Marques *et al.* 2009), to account for false positives; we also need to convert the total number of boing detections across hydrophones to number of unique boings. Hence, the proposed estimator is now:

Equation 2 $$\hat{D}=\frac{n_c(1-\hat{c})}{\hat{m}\times T\times\hat{a}\times\hat{r}}$$

where $\hat{m}$ represents the mean number of detections across the 14 hydrophones for each unique boing. While we present here an animal density estimator, we note that at this time there is no reliable estimate of cue rate ($\hat{r}$) for the minke boing sound, and hence our focus is on the density of boings.

We obtain the different random components in Equation 2 from different data sets. We obtain n_c from the sum across hydrophones and time periods of the counts of automatic detections for all available data (detected boing data set). The other components are obtained from the smaller sampled time periods, where data was manually associated (associated boing data), as described below. Because of this, it is convenient to rewrite the density estimator in equation 2 as

Equation 3 $$\hat{D}=\frac{\hat{\varphi}\times n_c}{\hat{r}\times T}$$

where the 3 random components estimated from the associated data are collapsed to a single parameter $\hat{\varphi}=\frac{(1-\hat{c})}{\hat{m}\times\hat{a}}$. Note (1-$\hat{c}$)/$\hat{m}$ converts the number of boings detections across all hydrophones to the number of unique boings detected.

Here, long periods of automatically detected counts were available, but manually associated data, which were much more labor intensive to obtain (see below), were restricted to shorter sampled periods of time. Hence the estimation approach taken was chosen to allow the use of all the data available. Nonetheless, density estimation could be based on only the manually associated boing data set. For comparison purposes, we also calculated density considering only the data from these shorter time periods. Note that in that case, estimation is based on a much simpler estimator, namely the usual conventional estimator in SECR

Equation 4 $\hat{D} = \frac{n}{\hat{a} \times \hat{r} \times T}$

where *n* represent the number of unique boings detected, pooled across time periods. This density and the corresponding variance are a direct output of the *secr* package.

Using the manually associated boing data set, we estimated both (1) the false positive proportion ($\hat{c}$), and (2) the multiplier $\hat{m}$. The false positive proportion was estimated as the weighted average of the false positive proportion in each of the twelve sample files which comprise the associated boing data set, with the weights being the number of boing detections (including false positives) in each file. On the other hand, $\hat{m}$ was estimated as the weighted average of number of detections across the 14 hydrophones for each unique boing produced, with the weights being the number of manually confirmed boing detections in the associated boing data set.

To convert from boing density to animal density we need an estimate of boing cue rate ($\hat{r}$). While we can really only report with an accepted level of confidence the boing density of minke whale boing calls, we also provide a preliminary boing calling minke whale density, using an overtly preliminary and unreliable cue rate, for the sake of illustrating the methods. The estimate is based upon the concept of tracking an individual minke whale that remains within the hydrophone array detection and localization area over a long time period which includes at least one quiet period. Counting the number of boings produced by the individual and dividing by the time period provides the preliminary boing rate. Key to this method is having confidence that one is detecting the same individual over the entire time period and that all boings produced by the individual are detected. Here we use multiple characteristics of the boing itself (fine detail of spectral characteristics and inter-boing-interval) along with spatial location and trend direction of the animal over the analysis time frame. This method was developed after observing in post analysis the high degree of stability of the frequency of the dominant spectral component (DSC) of boings for an animal sighted on 27 April 2009 at PMRF. The DSC is a refinement on the mode frequency (described in the boing signal processing section) which is used to help confirm that boing sound detections from multiple hydrophones are from a single boing call. We hypothesize the DSC has potential for helping identify individuals in cases such as this, where boings from one individual have DSC frequencies well separated from other individuals present nearby. This sighting was cued both by the PMRF hydrophone-determined location information that was radioed from shore, and by the presence of seabirds in the area of the sighting. It is our belief that the sighted animal can be attributed to producing a known number of boing calls in two periods of time separated by one quiet period, in over 12 hours of data making it possible to determine a preliminary boing cue rate (see details in the Results section).

Considering equation 3, assuming independence of the random components, we can approximate the variance of this product using the delta method (as in Marques *et al.* 2009):

Equation 5 $\mathrm{var}(\widehat{D}) \approx \hat{D}^2 \left\{ CV^2(n_c) + CV^2(\hat{\varphi}) + CV^2(\hat{r}) \right\}$

where CV(a) represents the coefficient of variation of the corresponding random component(a).

The variance in the number of counts is obtained by using a distributional assumption, namely the Poisson assumption, which is consistent with the usual assumption made in SECR studies (see *e.g.*, Efford and Borchers unpublished supplementary material to Borchers and Efford 2008).

In order to obtain the variance for the second component ($\hat{\varphi}$), we implemented a non-parametric bootstrap, resampling the available time periods. Hence, for each of 999 times, a resample with replacement of 12 time periods was obtained, and $\hat{\varphi}$ computed from it, by calculating the three relevant random components as described above based on the associated boings data set. The variance in $\hat{\varphi}$ was estimated as the empirical variance of the 999 bootstrap estimates.

Here we have not considered a variance for the cue rate ($\hat{r}$); as stated above, that estimate is based on a preliminary data set and used for illustration only, and we abstain from reporting measures of precision for that component and for animal density to stress the fact that these are preliminary and ultimately unreliable.

We assume that the boing density estimate has a log-normal distribution to obtain confidence intervals, as described in Buckland *et al.* (2001).

Data Collection and Signal Processing

This study is based on data collected in 2006 from 14 deep water (3.5-4.7 km) seafloor-mounted hydrophones, which are part of PMRF located off the northwest coast of Kauai (Fig. 1). We consider data from twelve days, recorded opportunistically in 2006: 5, 21, and 25 February; 5, 9, 13, 17, 25, and 29 March; and 2, 6, and 18 April. A personal computer-based recorder equipped with a 32-channel simultaneous sampling analog to digital converter (ICS Ltd 645A) operating at a 96 kHz sample rate with 16 bit samples was utilized to record the hydrophones utilized in this study. Data was recorded continuously each day, from approximately 8 AM until 10PM, and stored in 10-minute files until the storage drive was filled. This provides a total time of approximately 167 hours for each of the hydrophones utilized in this case study. These 14 hydrophones have adequate frequency response (0.1-18kHz) to detect boing;

however the hydrophone sensitivities are unavailable. While these data are primarily from daylight hours, it has been reported that minke whale boings in Hawaiian waters show no significant diel variations (Oswald *et al.* 2011).

An automatic boing detector, based upon a frequency contour whistle detection process (Mellinger *et al.*, 2011) was utilized for detecting minke boing vocalizations. The detector was tuned to a specific frequency band (1350-1440 Hz) where the dominant spectral component of the signal, as observed in Hawaiian waters, is typically located (Fig. 2), and was verified as part of an optimization of the detector parameters. The algorithm works by tracking spectral peaks over time, grouping together peaks in successive time-slices in a spectrogram if the peaks are sufficiently near in frequency and form a smooth contour over time.

A real-time version of the Minke boing detector was implemented on the Marine Mammal Monitoring on Navy Ranges system (Morrissey *et al.* 2006) which was utilized to process the twelve day data set. Output was a file of times and other statistics associated with each sound automatically classified as a minke whale boing call. We refer to this data set as the “detected boings” data set. This includes false positives (detections of sounds not boings and multiple detections of the same boing on a single hydrophone), which must be accounted for when estimating density (otherwise density is overestimated).

Association of the same boing sound detected on different hydrophones is required for the SECR analysis as well as an estimate of the false positive proportion; however the process can tolerate missed detections. Two hours of sample data was selected for further processing by chronologically concatenating all available 10-minute data files and selecting twelve of the10-minute samples systematically spaced with a random start. The aims of processing these data were twofold: firstly to associate boings detected on multiple hydrophones that came from the same original vocalization (*i.e.*, akin to a capture history in mark-recapture), and secondly to estimate the false positive proportion. The final product of the association process is referred to as the "associated data set".

The association process began with an automated algorithm, followed by manual validation. The automated association process operated over the detection reports in a temporally sequential manner, utilizing a sliding window of 28 seconds, which represents the maximum travel time of a unique boing between the farthest spaced hydrophones. When a detection is encountered on a sensor, detections with the same mode frequency (+/- our frequency measurement resolution of one FFT bin width, or 5.86 Hz) in the 28-second window were evaluated across all sensors. Detected boings which meet the time and frequency criteria were grouped into a unique boing associated group. For each associated group, the hydrophone number, detection time and mode of the frequency peaks were saved, with the unique boing number incrementing for subsequent associations. In cases where the same boing was detected multiple times as determined by multiple detections with the same mode frequency in a six second time period to account for a potential bottom reflected multipath, only the first detection was saved. Manual validation was performed by a single experienced analyst (SJ), using visual plots of spectrograms, spatial and temporal detection arrival times and aural monitoring to confirm that associations were valid, and that no false positive detections remained in the associated boings data set.

In the context of our density estimation method, false positive detections are defined as detections that are not minke whale boings and multiple detections of the same boing on a single hydrophone. The number of false positive detections in each 10-minute sample is the total number of detections in the sample minus the number of detections which were manually validated.

Sensitivity of density estimate to human manual boing association

We assume that manual associations are performed without error. This is probably optimistic, because there were instances in which slightly subjective decisions needed to be made. To assess the potential impact of these choices, the file from 17 March 2006, containing the largest number of detections (*i.e.*, the one in which manual association would be hardest) was independently processed by an additional operator (SWM), and the corresponding two density estimates for that time period compared. This allows an assessment of worst case scenario impacts on estimates due to human operator.

RESULTS

Estimating boing density and minke whale density

In the manually associated boings data set, over the twelve 10-minute periods, 204 unique boings were detected. From the SECR analysis, as for Marques *et al.* (2011*a*), the half-normal detection function model provided the most parsimonious fit to the capture histories obtained from the manually associated data, with the second-best model being the cumulative lognormal (ΔAIC=19.6), and with distant third- and fourth-best models being respectively the hazard rate (ΔAIC=93.7) and the negative exponential (ΔAIC=208.0). For all 4 models the inclusion of hydrophone depth as a covariate was not useful in explaining the scale parameter of the detection function (ΔAIC>1.9). Figure 3 depicts the estimated detection function as a function of distance. The estimated effective surveyed area of the array ($\hat{a}$) of the 14 hydrophones was 8,767 km^2.

The associated boing data set also provided the estimated false positive proportion, $\hat{c}$, of 0.194 and the estimated $\hat{m}$ multiplier of 3.899.

In the detected boing data set over the approximate 167 hours of the recordings, a total of 92,143 sounds were automatically detected and classified as minke whale boings (n_c).

Combining the parameters obtained from the detected boing and associated boing data sets (equation 3), we arrive at a density estimate of 130 boings per hour per 10,000 km^2, with an estimated CV of 11.5%, and the corresponding 95% CI being 103-163.

Considering the associated boing data set only – *i.e.*, using the standard SECR method (equation 4), we estimate density to be 116 boings per hour per

10,000 km^2, with an estimated CV of 10.6%, and the corresponding 95% CI being 95-143.

We explicitly report a minke whale density estimate (for animals producing the boing call) to draw the reader's attention to the fact that the cue rate is required in order to convert the cue density into a density of calling whales. While the long term boing cue rate for minke whales is currently unknown, here we present a very tentative cue rate based on a single data set for what is believed to be the minke whale sighted at 1350 HST on 27 April 2009 at PMRF. Boings were detected and localized in near real-time from shore between 1000 and 1344 with estimated position updates radioed to a field crew resulting in a sighting within a few hundred m of the last reported localization. Post-fieldwork analysis of recorded data indicates 57 boings suspected to be produced by the sighted individual over the 5.9h period between 0749 and 1350. The boings are attributed to an individual for the following reasons: successive localizations of these boings over this period are within a few hundred meters of one another with the last close to where the sighting occurred and no other boings were localized within ten km of this area over this time; the inter-boing intervals (mean IBI 377.4 s, SD=111 s) fit with the slower boing rate observed for minke boings in Hawaii (*e.g.,* Thompson and Friedl 1982); the dead reckoning course (321 degrees true) and rate of advance (mean 1.86 km/h) are consistent with that of an individual whale; the DSC frequencies are consistent (mean DSC 1,384.0 Hz, SD=1.78 Hz) and none of the thousands of other automatically detected boings over this period are within +/-15 Hz of 1,384 Hz. No other boings were detected near 1384 Hz for the next 4.7 h; an additional 18 boings were then detected at this frequency (mean DSC 1,384.1 Hz, SD=1.72 Hz; mean IBI=395.6 s, SD=188 s). The dead reckoning course of the source of these boings during the quiet period was 320 degrees true with a mean speed 3.8 km/h, which, while over double the previous rate, is reasonable for a minke whale and below that reported for minke in Hawaiian waters of 5.6 to 5.7 km/h (Rankin and Barlow 2005). This analysis was truncated at 2013 HST as the source of boings was heading out of the localization range of the array. If one makes the assumption that the boings at this frequency over this time are indeed attributed to a single individual, the 75 boings in a 12.41 h period result in a tentative average boing rate of 6.04 boings/h. Note that this boing rate represents only one sample for one half of a day and does not include any examples of faster boing rates.

Given this preliminary estimated boing rate, we can convert the above boing call density estimates to boing calling minke whale density estimates of 21.5 and 19.2 whales per 10,000 km^2, respectively, when considering the analysis based on the 167 h detected boings data set and the smaller sampled associated boings data set only.

Sensitivity of density estimate to human manual boing association

Based on a single 10-minute file from the associated data set the two different human operators had nearly identical overall boing counts (55 *vs*. 56). However the assignment of the automatic boing detections across hydrophones to unique boing calls were different enough that the density estimates obtained by SWM was 11.9% lower than that of the main operator (SJ). Note it is likely that both operators made minor mistakes in association.

DISCUSSION

Density estimates

The estimates obtained represent another example of animal density estimate based on acoustic data, a field that is currently in its infancy but which shows an enormous potential. Here we present a more rigorous boing call density estimate for Hawaiian waters as compared to the exploratory study in Marques *et al.* (2011*a*), and for illustration a tentative calling minke whale density estimate based on a very preliminary cue rate. This is especially relevant as visual observations have been rare in this geographic area. Perhaps surprisingly, the differences in density estimates obtained between using only the much smaller (2 h) associated boings data set and the considerably larger (167 h) detected boings data set were minor; even more surprising, the boing density estimate's precision obtained was very similar. This may be explained because these correspond not only to two different data sets, but two different estimation approaches: while the longer data set naturally contains more information, its use is made at the cost of estimating two extra parameters, namely $\hat{m}$ and $\hat{c}$. This seems to suggest that for future surveys aimed at obtaining acoustic-based density estimates such as ours, it might be more efficient to consider more sampled time periods of manually associated data than longer periods of automatically detected data. While here we considered only 12 time periods, a larger number of time periods should be used in future, to adequately sample the time over which inferences are desired. In fact, in our associated boings data set, 2 out of the 12 days considered had significantly more detections, potentially representing either higher densities of animals or periods of rapid boing rates. This large variance in the number of detections over time intuitively means that a larger number of time periods would be advisable to gain reliable mean estimates. Note that the overall result for the mean minke boing density of 130 boings per hour per 10,000 km^2 is over twice that reported in Marques *et al.* (2011*a*) of 48 boings per hour per 10,000 km^2. However, those results used half as much associated data as the current study (one hour *vs*. this study two hours), one third of the data was from a different year (2007), and the 10-minute periods used were chosen non-randomly as they were originally selected to test the automated detector algorithm.

In terms of north Pacific minke whale densities reported in the literature, the average minke whale density estimate provided by Zerbini *et al.* (2006) was around 60 whales per 10,000km^2 for vessel based visual surveys using survey blocks in the Aleutian Islands and Alaskan waters. The visual survey counts whales of both sexes and of all ages and the whales behaviors are different from behavior at lower latitudes in winter months. The preliminary boing calling minke whale density from above could be in reasonable agreement with the minke whale density determined by Zerbini *et al.* in Alaskan waters if one conjectures that approximately one third of the population (*e.g.,* the reproductively active males) may be making the boing vocalization. It is clear that more research is needed in order to convert the minke call density into a density of whales.

Method assumptions

The SECR methods used have a number of explicit and implicit assumptions, namely that: (1) associations are made without error, (2) that the boing sounds are uniformly distributed in 2D (horizontal) space, (3) that the detection process is well modeled, namely that it is a function of distance to the hydrophone and boings produced over the hydrophones are detected with certainty (in this study; not a general feature of SECR); (4) that hydrophones detect boings independently. Further, to estimate animal density, (5) the cue rate must be an unbiased estimate of the (unknown) cue rate observed during the survey period. Finally, it must be noted that (6) the estimate is valid only for the fraction of the population which is actively producing boings. We address these in turn below.

Regarding the reliability of manual associations, the comparison of results from using manual associations from different human operators is reassuring. Given we deliberately chose the worst case scenario to quantify this potential problem, differences in density below 15% represent a reasonable upper bound to the potential impacts. We anticipate much smaller differences when less active periods are used to estimate the effective survey area, which could in fact be purposefully chosen if detectability can be safely assumed independent of density.

The true distribution of boings in space is unknown, and the uniform distribution is essentially a working hypothesis which seems a reasonable approximation lacking any better model. We note that other models could be implemented, namely an inhomogeneous Poisson process (*e.g.*, Borchers and Efford 2008). It is nonetheless reassuring that SECR methods have been reported to be relative insensitive to violation of this assumption (*e.g.*, Efford *et al.* 2009). We purposely have not used a small number of hydrophones available closer to the islands because the number of boing detections on these was far lower than that observed in the hydrophones used here (Steve Martin, unpublished results). This could be an issue regarding different detectability or different availability for detection, but due to the scale of the problem it seemed more likely a detection problem and we avoided addressing it here. This might be the object of future research.

Regarding the detection process, although we have not presented the results here, we note that models for which some sounds were not necessarily detected if produced directly above the hydrophone (*i.e.* at horizontal distance 0) seemed, according to AIC, to provide fits as good or even better than the half normal model used here. This might deserve further investigation, as it could be a hint that of some unexpected behavior of the acoustic system used. Nonetheless, and reassuringly, we note that the density estimates were very insensitive to the detection model used, even for models for which the intercept was not 1. This is consistent with Efford *et al.* (2009, and references therein), who noted the robustness of the density estimates to the detection function used, and makes this a point of lesser concern in practice. Nonetheless, the development of methods to check goodness-of fit of detection models in SECR analyses is an important area of future research. The detection function we estimated by SECR seem reasonable in that the probability of detecting a boing at 10km horizontal distance is in the range of 0.8 to 0.9. Minke boing vocalizations are typically detected as direct path arrivals out to approximately 25 km, with bottom-surface-bottom multipath arrivals detectable well past the direct path distance. The source level of the minke boing is currently only estimated and the hydrophones utilized in this study do not have calibration data. The depth of the vocalizing minke whales is unknown, but assumed to be within the upper couple of hundred meters of depth. The detector utilized in this study processes a limited frequency range (1,350 to 1,440 Hz) which has favorable absorption loss (on the order of 0.1 dB per km). Making assumptions of whale depth of 100m while vocalizing in waters of 4.1 km deep with a source level of 150 dB re micro Pascal (uPa), the expected direct path signal level received at the bottom hydrophone 10km distant would be on the order of 69 dB assuming spherical spreading and accounting for absorption loss. This signal level should be detectable for deep water ambient noise for moderate sea states (using 45dB re uPa^2Hz^{-1} spectrum noise level). The assumption that boings produced directly over the hydrophone are detected with certainty is not a SECR requirement. However, using the above assumptions the signal level would be approximately 77 dB which should be detected with near certainty even though the hydrophone is still 4km distant from the source in the depth dimension.

The independence of detections across hydrophones seems reasonable as their operation and the data processing from each one are completely separate processes.

Obtaining an adequate cue rate to convert sound density to animal density is a fundamental step in all cue based density estimation methods. As reported by Marques *et al.* (2011*b*), the cue rate might be by far the largest contributor to the overall density variance estimate. The apparent bimodal cue rate for minke boing calls complicates obtaining a reliable long term cue rate. In addition, it is not certain which minke whales produce the boing call, nor the ratio of boing calling whales to total whales. We avoided giving emphasis to our animal density estimate because it should be viewed with extreme caution, being based on one small and preliminary data set which is believed to represent a single whale for one 12 h period with only the slower cue rate present. The method presented for obtaining boing cue rate by acoustically identifying individuals using the DSC is an interesting hypothesis and appears valid for the single case presented on 27 April 2009, where only a couple of animals were acoustically detected on the range with large separation both spatially and in their DSC frequencies. Without having data from acoustic tags directly measuring individuals' acoustics it is difficult to prove the DSC method is identifying individual whales. The method struggles in cases of high density of animals, where animals with similar DSC frequencies are located close to one another in the study area. Given the current active development of passive acoustics density based methods, we anticipate considerable research focused at estimating cue rates, in particular describing relationships between cue rate and relevant covariates, such as season, animal behavioral state, density, and proximity of ships.

If the boing call is a mating display produced only by reproductively active males, not only there is a potential for cue rate density dependence, there is a proportion of the population which cannot be detected using our acoustic methods. Naturally the density estimate reported here corresponds to the fraction of the population which is producing the boing sounds during the survey period. If a random sample of animals can be obtained not depending on acoustics (which might be complicated in our particular setting but feasible if many animals are tagged with acoustic tags) the proportion of the silent animals might be automatically accounted for if the adequate cue rate is obtained. In fact, provided a random sample of animals is used to estimate cue rates (and hence the proportion of silent animals is representative of that observed in the population, and their data included as zero cues per unit time), the resulting cue rate, for silent and acoustically active animals combined, would lead to a density valid for the entire population, not only for the sound producing animals.

It might seem strange that we used a habitat mask spanning over 200 km from the hydrophones, but this should not be over-interpreted. Such a large distance is used just for caution, as the only shortcoming in using a larger distance is the computational burden. In fact, the same results for the half normal model would have been obtained if an 80 km buffer distance was used. However, because some of the models used, like the negative exponential and the hazard rate, have heavy tails (*i.e.*, a considerable amount of the distribution is in the long tail), we opted to use a larger distance. Additional investigation showed that a distance of over 1,000 km should be used for the negative exponential and hazard rate. This seems to imply that, due to the heavy tail, these are not plausible models for the detection function in the context of SECR.

The boing density could also be obtained in alternative ways. As an example, another approach could consider the density estimated by each hydrophone and average the hydrophone specific estimates, although different, perhaps less reliable assumptions, would be needed. The method described here considered a common false positive rate across all hydrophones. Another option would be to consider hydrophone-specific rates. This presents no additional requirements apart from the need for much more intensive sampling (as rather than estimating an average of 14 values, it requires the estimation of 14 independent values) and the consequent human operator time.

Conclusion

Our minke whale boing density estimate provides density information for a species in Hawaiian waters where significant aerial and shipboard visual survey efforts have not produced density estimates due to insufficient sightings.

Acoustic based methods are likely to become widespread in the future, and much research is anticipated in this area, regarding cue rates, acoustic behavior, and animal sound processing hardware and software. Where arrays of sensors exist across which association of sounds can be made, leading naturally to capture histories amenable to capture recapture methods, SECR methods seem to be a strong candidate to obtain density estimates. These techniques are envisioned to enable future systematic estimation of species densities at areas of dense hydrophone arrays.

The use of a mobile array of 4 sensors moved throughout a survey area (Dawson and Efford 2009) illustrates methods in which small mobile arrays can be employed in data collection for using SECR methods to estimate density. This concept can also be applied to the marine environment and may be more practical for areas where large arrays of hydrophones do not exist.

ACKNOWLEDGEMENTS

We appreciate the support of personnel at the Pacific Missile Range Facility, notably Jim Hager, Mike Dick, Eliseo Bolosan and Robin Higuchi, in collection of acoustic data over several years. This work is an output of the National Oceanographic Partnership Program sponsored project Density Estimation of Cetaceans from passive Acoustic Fixed sensors (DECAF), funded jointly by the Joint Industry Program of the International Association of Oil and Gas Producers and the US National Marine Fisheries Service. TAM research partially sponsored by national funds through the Fundação Nacional para a Ciência e Tecnologia, Portugal – FCT under the project (Pest OE/MAT/UI0006/2011). We thank all the DECAF team members for discussions that led to a better paper, particularly David Borchers who provided insightful comments regarding the SECR analysis. The data used in this study are freely available on the Ocean Biogeographic Information System, Spatial Ecological Analysis of Megaverebrate Populations (OBIS-SEAMAP) web site.

LITERATURE CITED

Akaike, H. 1985. Prediction and entropy. Pages 1-24 *in* Atkinson, A.C. and Fienberg, S.E., Eds. A celebration of statistics: the ISI centenary volume. Springer, New York.

Balcomb, K.M. 1987. The whales of Hawaii, including all species of marine mammals in Hawaiian and adjacent waters. Marine Mammal Fund, San Francisco, CA.

Borchers, D., and M. Efford. 2008. Spatially explicit maximum likelihood methods for capture-recapture studies. Biometrics *64*:377-385.

Buckland, S., D.R. Anderson, K.P. Burnhanm, J.L. Laake, D.L. Borchers, and L. Thomas. 2001. Introduction to distance sampling. Oxford University Press.

Carretta, J.V., K.A. Forney, M.M. Muto, J. Barlow, J. Baker, B. Hanson, and M.S. Lowry. 2005. U.S. Pacific Marine Mammals Stock Assessments. NOAA Technical Memorandum NMFS. NOAA-TM-NMFS-SWFSC-375. 316 pp.

Dawson, D.K., and M.G. Efford. 2009. Bird population density estimated from acoustic signals. Journal of Applied Ecology 46:1201-1209.

de Boer, M. 2010. Spring distribution and density of minke whale *Balaenoptera acutorostrata* along an offshore bank in the central North Sea. Marine Ecology Progress Series 408:265-274.

Efford, M. G. 2004. Density estimation in live-trapping studies. Oikos 106: 598-610.

Efford, M.G. 2009. *secr* – Spatially explicit capture recapture in R, version 1.2.10.1. Department of Zoology, University of Otago, Dunedin, New Zealand.

Efford, M.G., D.L. Borchers, and A.E. Byrom. 2009. Density estimation by spatially explicit capture-recapture: likelihood-based methods. Pages 255-269 *in* D.L. Thomson, E.G. Cooch and M.J. Conroy, eds. Modeling demographic processes in marked populations. Springer New York.

Gedamke, J., D.P. Costa, and A. Dunstan. 2001. Localization and visual verification of a complex minke whale vocalization. Journal of the Acoustical Society of America 109:3038-3047.

Johnson, M.P., and P.L.Tyack. 2003. A digital acoustic recording tag for measuring the response of wild marine mammals to sound. IEEE Journal of Oceanic Engineering. 28: 3-12.

Küsel, E.T., D.K. Mellinger, L. Thomas, T.A. Marques, D.J. Moretti, and J. Ward. 2011. Cetacean population density estimation from single fixed sensors using passive acoustics. Journal of the Acoustical Society of America 129: 3610-3622.

Marques, T.A., L. Thomas, J. Ward, N. DiMarzio, and P.L. Tyack. 2009. Estimating cetacean population density using fixed passive acoustic sensors: an example with Blainville's beaked whales. Journal of the Acoustical Society of America 125:1982-1994.

Marques, T.A., Thomas, L., Martin, *et. al.* 2011*a*. Spatially explicit capture recapture methods to estimate minke whale abundance from data collected at bottom mounted hydrophones. Journal of Ornithology. DOI: 10.1007/s10336-010-0535-7

Marques, T. A., L. Munger, L. Thomas, S. Wiggins, and J.A. Hildebrand. 2011*b*. Estimating North Pacific right whale (*Eubalaena japonica*) density using passive acoustic cue counting. Endangered Species Research. 13:163-172.

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

Moretti, D., Marques, T.A., Thomas, *et al.* 2010. A dive counting density estimation method for Blainville's beaked whale (*Mesoplodon densirostris*) using a bottom-mounted hydrophone field as applied to a Mid-Frequency Active (MFA) sonar operation. Applied Acoustics 71: 1036-2042.

Morrissey, R.P., J. Ward, N. DiMarzio, S. Jarvis, and D. Moretti. 2006. Passive acoustic detection and localization of sperm whales (*Physter macrocephalus*) in the tongue of the ocean. Applied Acoustics 67:1091-1105.

Oswald, J.N., W.W.L. Au, and F. Duennebier. 2011. Minke whale (*Balaenoptera acutorostrata*) boings detected at the Station ALOHA Cabled Observatory, Journal of the Acoustical Society of America 129: 3353-3360.

R Development Core Team (2009). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Rankin, S., and J. Barlow. 2005. Source of the North Pacific ‘boing’ sound attributed to minke whales. Journal of the Acoustical Society of America 118:3346-3351.

Rankin, S., T.F. Norris, M. Smultea, C. Oedekoven, A. Zoidis, E. Silva, and J. Rivers. 2007. A visual sighting and acoustic detections of minke whales, *Balaenoptera acutorostrata* (Cetacea: Balaenopteridae), in nearshore Hawaiian waters. Pacific Science 61:395-398.

Reeves, R.R., B.S. Stewart, P.J. Clapham, and J.A. Powell. 2002. Guide to Marine Mammals of the World. Alfred A. Knopf, New York: 525 pp.

Royle, J. A. 2009. Analysis of capture-recapture models with individual covariates using data augmentation. Biometrics 65:267-274.

Stewart, F.S. and S. Leatherwood. 1985. Minke whale *Balaenoptera acutorostrata* Lacepede, 1804 Pages 91-136 *in* Ridgway, S.H. and Sir R. Harrison Eds, Handbook of marine mammals Volume 3 Academic Press.

Thompson, P.O., and W.A. Friedl. 1982. A long term study of low frequency sounds from several species of whales off Oahu, Hawaii. Cetology 45:1-19.

Ward, J.A., Thomas, L., Jarvis, et al. in press. Passive acoustic density estimation of sperm whales in the tongue of the ocean, Bahamas. Marine Mammal Science.

Wentz, G.M. 1964. Curious noises and the sonic environment in the ocean. Pages 101-119 in W.N. Tavolga, Ed. Marine bioacoustics Vol. 1. Pergamon, New York.

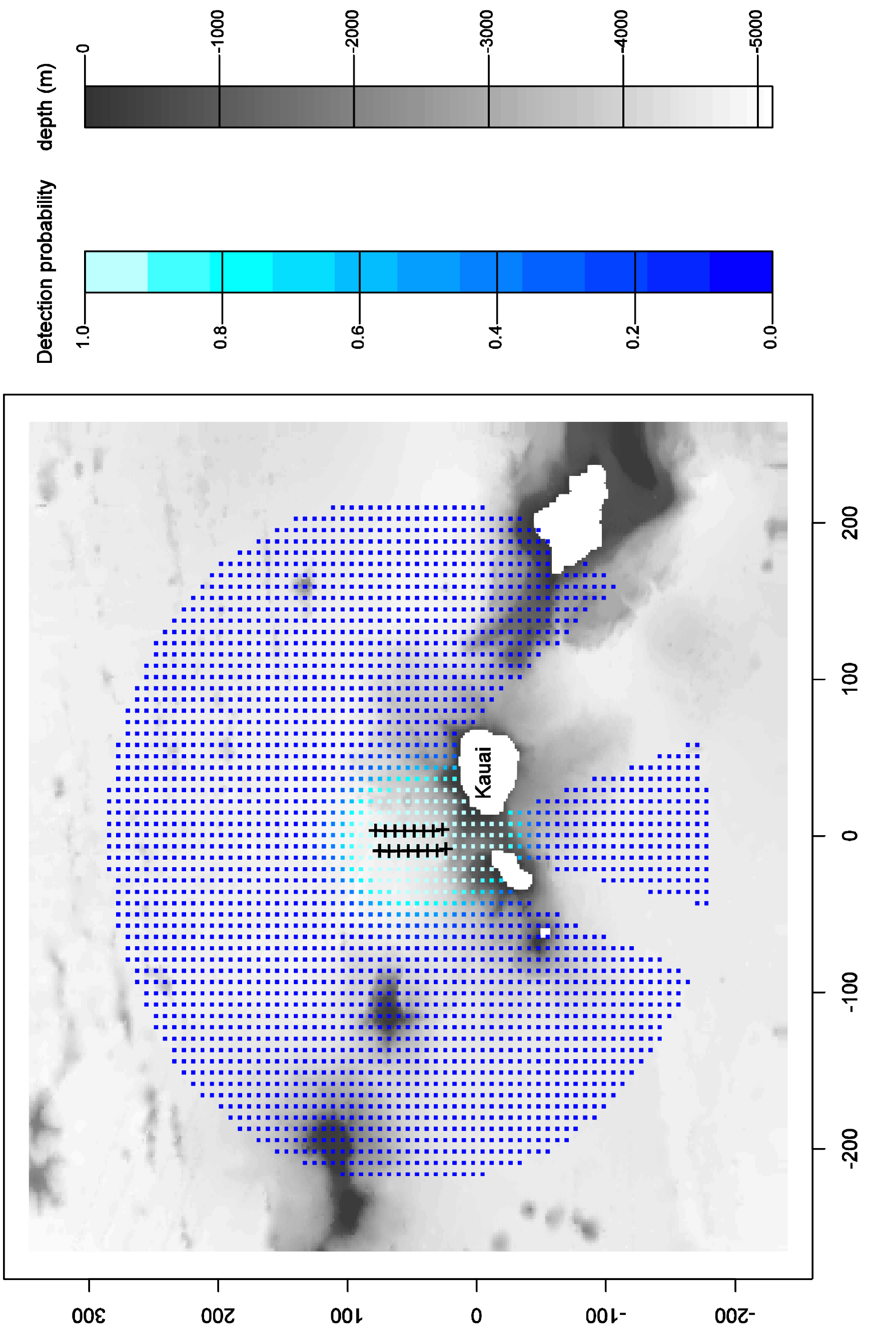
Zerbini, A.N., J.M. Waite, J.L Laake, and P.R. Wade. 2006. Abundance, trends and distribution of baleen whales off Western Alaska and the central Aleutian Islands. Deep-Sea Research I 53: 1772-1790.

FIGURE CAPTIONS

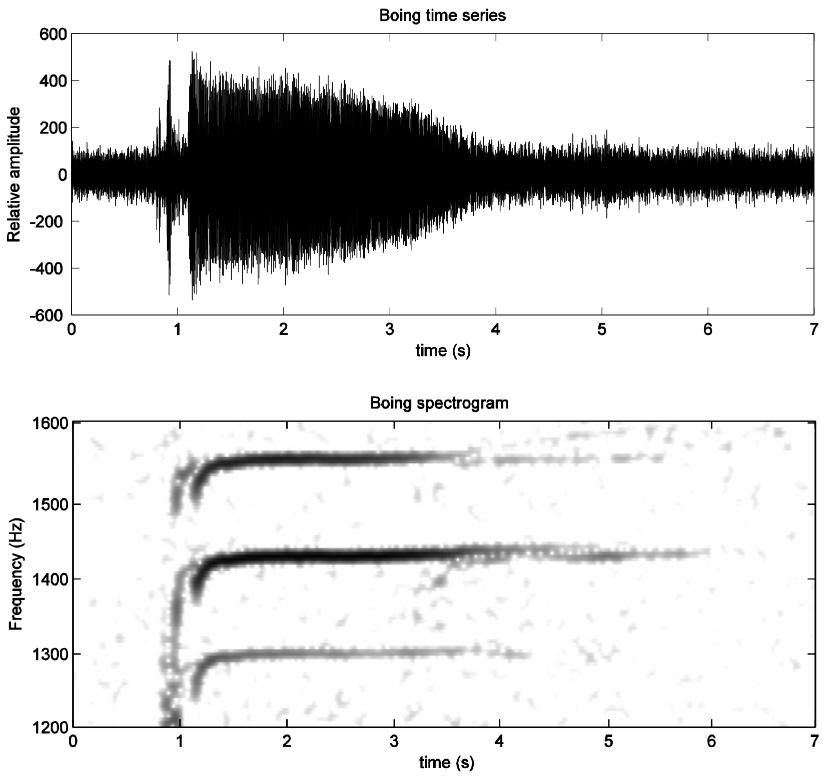
Figure 1 – Study area showing hydrophone locations (black crosses) with habitat mask represented by small square dots (color, online, proportional to estimated detection probability at each grid point). Island masses are represented by white polygons, and acoustic shadow zones have no dots. Approximate bottom depth is shown in gray scale. Horizontal and vertical scale units in km.

Figure 2 – Example of a minke boing vocalization time series and spectrogram.

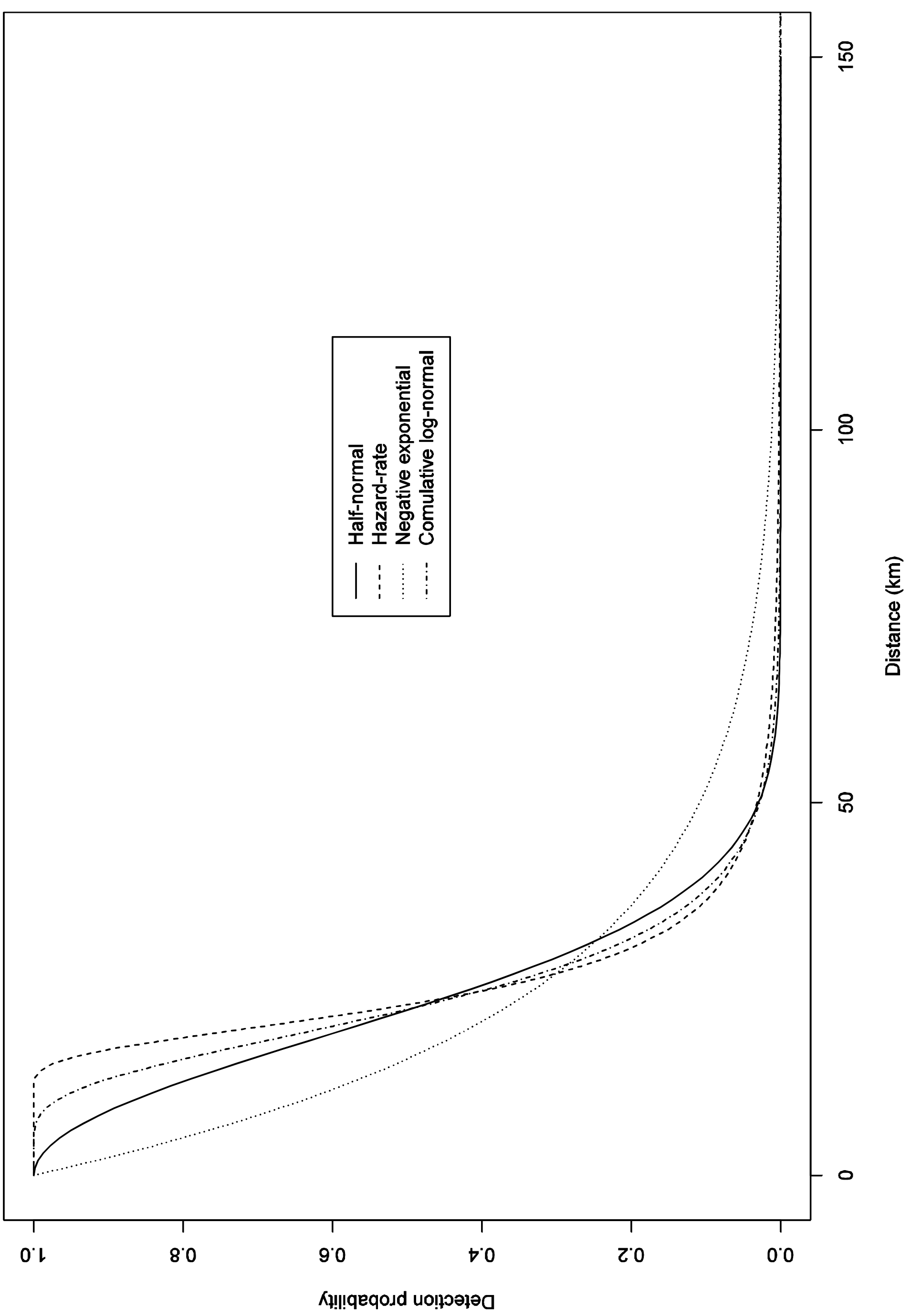
Figure 3 – Probability of detecting a boing as a function of distance for the different detection functions fitted. Model used for further inference is half-normal (solid line).



Formatted: Font: (Default) Arial



Formatted: Font: (Default) Arial



Formatted: Font: (Default) Arial