Cetacean population density estimation from single fixed sensors using passive acoustics

Elizabeth T. Küse(a) and David K. Mellinger
Cooperative Institute for Marine Resources Studies (CIMRS), Oregon State University, Hatfield Marine Science Center, Newport, Oregon 97365

Len Thomas
Centre for Research into Ecological and Environmental Modelling, University of St. Andrews, St. Andrews KY16 9LZ, Scotland

Tiago A. Marques
Centro de Estatı́stica e Aplicações da Universidade de Lisboa, Campo Grande, 1749-016 Lisboa, Portugal

David Moretti and Jessica Ward
Naval Undersea Warfare Center Division, Newport, Rhode Island 02841

(Received 17 December 2010; revised 18 March 2011; accepted 30 March 2011)

Passive acoustic methods are increasingly being used to estimate animal population density. Most density estimation methods are based on estimates of the probability of detecting calls as functions of distance. Typically these are obtained using receivers capable of localizing calls or from studies of tagged animals. However, both approaches are expensive to implement. The approach described here uses a MonteCarlo model to estimate the probability of detecting calls from single sensors. The passive sonar equation is used to predict signal-to-noise ratios (SNRs) of received clicks, which are then combined with a detector characterization that predicts probability of detection as a function of SNR. Input distributions for source level, beam pattern, and whale depth are obtained from the literature. Acoustic propagation modeling is used to estimate transmission loss. Other inputs for density estimation are call rate, obtained from the literature, and false positive rate, obtained from manual analysis of a data sample. The method is applied to estimate density of Blainville’s beaked whales over a 6-day period around a single hydrophone located in the Tongue of the Ocean, Bahamas. Results are consistent with those from previous analyses, which use additional tag data.


PACS number(s): 43.30.Sf, 43.80.Ka [WWA] Pages: 3610–3622

I. INTRODUCTION

Passive acoustic methods can be valuable tools for monitoring populations of marine mammals since many species produce loud and distinctive calls. Acoustic recordings of marine mammal vocalizations, including echolocation clicks, calls, and songs, sometimes used in conjunction with line-transect methods (e.g., Barlow and Taylor, 2005), are being used increasingly often to better understand a particular species’ distribution and population density in a given geographic area (McDonald and Fox, 1999; Moretti et al., 2006; Mellinger et al., 2006, 2007; Marques et al., 2009). Recently, Marques et al. (2009) presented a density estimation method based on counting the number of detected cues, where a cue is an effectively instantaneous event such as a single vocalization or the start of a vocalization sequence. The method also requires information on the probability of detecting a cue as a function of distance, the rate at which animals produce cues, and the fraction of false detections. This method was applied to estimate the density of Blainville’s beaked whales (Mesoplodon densirostris) over a 6-day period in 2005 at the Atlantic Undersea Test and Evaluation Center (AUTEC) in the Tongue of the Ocean, Bahamas. Results of their analysis yielded an estimate of 25.3 [95% confidence interval (CI) 17.3–36.9] or 22.5 (95% CI 15.4–32.9) animals per 1000 km², depending on assumptions about false positive detections.

Estimating the detection function, i.e., the probability of detecting a cue as a function of distance from a receiving sensor (Buckland et al., 2001), is a key element required by the above density estimation method. The estimated detection function is integrated over the area of interest, which is equivalent to averaging over distance, assuming that animals are uniformly distributed around the sensors, to derive the estimated average probability of detection. One method for calculating the detection function from data recorded at widely spaced hydrophones in an array configuration is to use the time difference between arrivals of direct path signals at the sensors to estimate the positions of vocalizing animals, and hence, their distances with respect to the sensors. In their study, Marques et al. (2009) used data from 82 hydrophones from the AUTEC array. However, due to the narrow beam pattern of Blainville’s beaked whales and the configuration of the hydrophones at AUTEC, clicks are normally detected at only one or two hydrophones simultaneously (Ward et al., 2008). Such conditions make methods based on localization...
difficult. One recently developed alternative (Marques et al., 2010) makes use of multiple detections without precise time differences, and hence bypasses the need for localization. However, a reasonable number of multiple detections are required, and these are not obtained in the case of beaked whales. Consequently, to estimate the detection function for Blainville’s beaked whales, Marques et al. (2009) made use of auxiliary data sets recorded with digital acoustic tags (DTAGs) (Johnson and Tyack, 2003) at AUTEC. Emitted clicks were detected both in data from individual tagged animals and in data from surrounding hydrophones. These detections were combined to localize precisely the animals using a hyperbolic fixing algorithm. The number of clicks detected on each hydrophone was measured and compared to the number produced by the animal to derive the probability of detection as a function of distance. Another important variable required for density estimation, the click production rate, was also estimated using DTAG data.

The goal of this work is to develop and demonstrate an alternative method of estimating animal density from acoustic recordings on separate, non-linked fixed hydrophones. By using separate, non-linked sensors, the method does not rely on localizing animals, a process that can be labor intensive even when it is feasible. Necessary information on animal vocal behavior is preferably obtained from the literature or derived from known information on similar species, instead of using acoustic tag data. Affixing acoustic tags to marine mammals is both difficult and expensive, and most often tag data from a specific population are not available. However, it should be noted that information available in the literature for certain species such as beaked whales, may come mostly from acoustic tag studies, although not necessarily from the same population or area being considered.

The method developed here is applied to estimate the density of Blainville’s beaked whales at AUTEC during the same time period analyzed by Marques et al. (2009), and the results are compared. The current study focuses on data from a single hydrophone and follows the methodology outlined by Zimmer et al. (2008) to estimate the detection function by application of the passive sonar equation. The probability of detecting a click as a function of distance from a single sensor is estimated by simulating clicks at a set of distances, predicting their signal-to-noise ratio (SNR) at the sensor using the sonar equation, and determining whether they would be detected at this SNR using information about the sensitivity of the detection and classification system. The use of the sonar equation requires knowledge of animal source level and characteristics of their sound beam pattern, including change in source level as a function of off-axis angle, propagation loss from source (whale) to receiver, and ambient noise levels at the receiver. Distributions of these parameters are used in a Monte Carlo algorithm to simulate the probability of detecting beaked whale clicks. The animal’s location and orientation relative to the sensor must be simulated in order to estimate the acoustic off-axis angle and hence the attenuation of source level. The slant distance to the sensor is also important to infer sound transmission losses, which occur due to signal spreading and sound absorption. The on-axis source level, animal depth and orientation, and click production rate are gathered from information available in the literature derived from acoustic tags, but not the same acoustic tag data sets used by Marques et al. (2009). The sensitivity of the detector is characterized by manually analyzing a short section of data to determine the SNR of all clicks, then analyzing the same data using the detector to determine which clicks were correctly detected and classified, and finally using this to build a regression curve of probability of detection versus SNR. The estimated average probability of detection from the simulated clicks is then used to estimate density of whales at AUTEC using the density estimation formula from Marques et al. (2009).

II. BEAKED WHALES AT AUTEC

The family of beaked whales, Ziphiidae, is composed of 21 deep-diving species that spend little time at the surface and hence are difficult to observe and study (Barlow et al., 2006). Until about a decade ago, little information was available on their ecology, behavior, or vocalizations. Strandings of two species, Cuvier’s (Ziphius cavirostris) and Blainville’s beaked whales, coincident with U.S. Navy sonar operations, prompted many efforts to study these species in more detail to understand their distributions, their use of sound, and the effects of sound on them (Barlow and Gisiner, 2006).

This paper is focused on Blainville’s beaked whales, although references to Cuvier’s beaked whales are made occasionally since both species have somewhat similar acoustics (Tyack et al., 2006). These species produce high-frequency echolocation clicks with center frequency in the range of 30 to 40 kHz. They also produce two distinct types of clicks during foraging (Johnson et al., 2006). One type, called a regular click, is frequency-modulated; the other, the buzz click, is believed to occur during the final stages of prey capture. Regular clicks have duration of approximately 250 µs, while buzz clicks are distinctly different from regular clicks, having mean duration of 104 µs. Buzz clicks are excluded from this analysis, as has been done elsewhere (Zimmer et al., 2008; Marques et al., 2009), and henceforth “click” refers to a regular foraging click.

Blainville’s beaked whale data were obtained from AUTEC, which includes a U.S. Navy undersea tracking range with an extensive array of bottom-mounted hydrophones suspended approximately 4–5 m off the ocean bottom at depths of up to 2000 m. Recordings from AUTEC’s hydrophones made with a sampling rate of 96 kHz and 16-bit resolution during a 6-day period in the spring of 2005 (April 26 to May 2) are used as the principal data set for this study. This is also the primary data set used by Marques et al. (2009) in a multi-sensor analysis to estimate density of Blainville’s beaked whales at AUTEC. The hydrophone array configuration used for the study is shown in Fig. 1 and the single sensor chosen (H 57) for analysis is also indicated. The bandwidth (±3 dB) of the single sensor is 50 Hz to 45 kHz, though some energy was present up to 48 kHz. The detection and classification system used to extract Blainville’s beaked whale clicks consisted of a multi-stage fast Fourier transform (FFT)-based energy detector. Detections of Blainville’s beaked whale clicks were classified by comparing in-band
(24–48 kHz) to out-of-band (0–24 kHz) energy (DiMarzio et al., 2008; Ward et al., 2008).

### III. DENSITY ESTIMATION

Marques et al. (2009) introduced a canonical formula for estimating the density of animals in a given area based on acoustic cue counting methods and on the formulation of Buckland et al. (2001). This formula is applied here, defining a cue as a foraging click, to estimate the spatial density of Blainville’s beaked whales in the vicinity of a single sensor.

The density estimation formula is given by

\[
\hat{D} = \frac{n_c(1 - \hat{c})}{K \pi w^2 \bar{P} \bar{r}},
\]

where \(\hat{D}\) is the density estimate, \(n_c\) is the number of detected clicks, \(\hat{c}\) is the estimated proportion of false positive detections, \(K\) is the number of sensors, \(w\) is the maximum detection distance beyond which we can safely assume that no clicks are detected, \(\bar{P}\) is the estimated average probability of detecting a click out to distance \(w\), \(T\) is the duration of the data set being analyzed, and \(\bar{r}\) is the estimated click production rate. Given variances on the input parameters, and assuming they are statistically independent, variance in estimated density can be obtained using the \(\delta\) method (Seber, 1982), and confidence limits can be obtained assuming the density estimate has a lognormal distribution (Marques et al., 2009). Unlike Marques et al. (2009), who used multiple hydrophones that gave a spatial sample of a large area, here we focus on the area around a single hydrophone. Therefore, while there is no variance component associated with spatial variation in \(n_c\), the variability in the number of detected clicks due to the randomness of the detection process needs to be accounted for. Here, counts were assumed to follow a Poisson distribution.

In the present analysis only data recorded on hydrophone 57 during the 6-day period are used and hence \(K\) in Eq. (1) equals 1. The total time of the recordings analyzed from the 6-day period \((T)\) is the same as from Marques et al. (2009), i.e., 4961 min. From this data set, the number of all clicks automatically detected and classified \((n_c)\) as Blainville’s beaked whale clicks on hydrophone 57 was 53 403. A maximum distance \((w)\) of 8 km is chosen to match that used by Marques et al. (2009), as detection distances of up to 6.5 km away from hydrophones at AUTEC have been reported for Blainville’s beaked whales (Ward et al., 2008). The estimation of the average probability of detecting a click \((\bar{P})\) and the click production rate \((\bar{r})\) are discussed in Secs. IV and V, respectively.

The proportion of false positive detections \((\hat{c})\) comes from comparison of manual and automatic detections (Marques et al., 2009). From the entire data set recorded during the 6-day period, 29 samples of 10 min each were used to estimate the false positive rate. The analysis of these samples yielded a total of 1958 sounds automatically detected on hydrophone 57, of which 1370 (almost 70%) were unambiguously identified as Blainville’s beaked whale clicks by a trained analyst. The proportion of true positives, \((1 - \hat{c})\), is estimated to be 0.703, with a standard error (SE) of 0.0418 and corresponding coefficient of variation (CV) of 5.94%.

### IV. ESTIMATING CLICK DETECTION PROBABILITY

In this section the estimation of the probability of detecting foraging clicks from Blainville’s beaked whales at a single fixed sensor is discussed. This is done by simulating animal sound production in an area around the sensor, then employing the passive sonar equation to estimate the SNR at the receiver. The estimated SNR is combined with the performance of the detection and classification system to estimate the probability that a given simulated vocalization is detected. Calculating a large number of these simulated vocalization events allows derivation of the detection function; calculating the average of this function over the area within distance \(w\) from a sensor results in the estimated average probability of detection \((\bar{P})\). Repeating this process many times, each time sampling from a distribution of each input variable, allows calculation of the standard error of \(\bar{P}\).

The SNR of a Blainville’s beaked whale click at the receiver can be estimated by application of the passive sonar equation (Urick, 1983), written in terms of acoustic intensity level \((I)\), as follows:

\[
\text{SNR} = \text{SL} - \text{DL} - \text{TL} - \text{NL},
\]

where \(\text{SL}\) is the whale’s on-axis source level (measured in dB re 1 \(\mu\)Pa at 1 m); \(\text{DL}\) is directivity loss, the attenuation (in dB) of the on-axis source level at a given angle from the animal’s acoustic axis; \(\text{TL}\) is sound transmission loss as a function of distance between source (whale) and receiver (in dB); and \(\text{NL}\) is the ambient noise level (measured in dB re 1 \(\mu\)Pa). Equation (2) implies that the acoustic power per unit area (or acoustic intensity) of a received signal is averaged over the duration of the signal \(\tau_r\). By considering the duration of the signal, a transient form of the passive sonar equation, written in terms of energy flux density \((E)\), can also be written and used to estimate the received SNRs (Urick, 1983;
Energy flux density relates to the acoustic intensity by

\[ I = \frac{E}{\tau}, \]  

(3)

Considering the above relationship, the source level (SL) of a transient signal is defined as

\[ SL = 10\log_{10}(E) - 10\log_{10}(\tau_i) = SLE - 10\log_{10}(\tau_i), \]  

(4)

where \( SLE \) is the source energy flux density (measured in dB re 1 \( \mu \)Pa² s), and hence, the energy-to-noise ratio (ENR) of a stationary signal is calculated as

\[ ENR = SLE - DL - TL - NL, \]  

(5)

which is the transient form of the passive sonar equation. Equations (2) and (5) are then related by (Au, 1993)

\[ ENR = SNR + 10\log_{10}(\tau_i). \]  

(6)

Each parameter in Eq. (5) is characterized by a distribution used in a Monte Carlo model described below. Equation (6) is then used to convert the simulated ENR into SNR levels, for which the probabilities of detection are estimated. The detection and classification system is characterized by the probability of detecting a single click as a function of SNR, so that simulated SNRs from the Monte Carlo model can be used to estimate the average probability of detection \( \hat{P} \). Note that the time window of measured and simulated SNRs must have the same duration.

To use Eq. (5) to estimate the SNR of received signals, the location of the sound source with respect to the receiver must also be simulated. Therefore, distributions of animal horizontal distances, clicking depths, and orientation angles with respect to the sensor were derived from the literature. These are required to calculate distributions of off-axis angles of an animal’s acoustic beam, which in turn is required to estimate the off-axis attenuation of source level (DL). Animal-sensor geometry is also important when using a propagation model to calculate acoustic transmission loss (TL). The assumptions made about animal position and orientation, source level, off-axis attenuation, transmission loss, and noise levels are presented in Secs. IV A–IV E. Section IV F discusses how the detector (and classifier) was characterized. The combination of these distributions into Eq. (5) using a Monte Carlo simulation is then presented in Sec. IV G.

A. Animal location and orientation

A first step in estimating the probability that a click will be detected at a given sensor is to simulate the position and orientation of the animal at the moment the click is produced. Detailed information on animal movement and vocalization behavior can be obtained from acoustic tag data. Here the goal is to estimate density without the use of local tag information, but information from the literature on both Blainville’s and Cuvier’s beaked whales is used to characterize the location and orientation of a whale with respect to a single hydrophone. It should be noted that the information found in the literature comes mostly from studies that utilized acoustic tags but that these studies were not done at AUTEC.

Discrete animal positions and orientations with respect to the sensor are derived from knowledge of the animal’s diving behavior during vocal periods. Position is estimated in three dimensions (\( x, y, \) and \( z \)) and orientation in two (azimuth \( \gamma \) and pitch \( \beta \), Fig. 2). The model used here assumes that beaked whales are randomly distributed in the horizontal plane (\( x, y \)) within a circle of radius \( w \) (defined earlier as \( w = 8 \) km) centered on the sensor. The random distribution in \( x \) and \( y \) implies a triangular distribution of distances from the sensor, with larger distances corresponding to linearly larger areas. Azimuth, or whale heading \( \gamma \) with respect to magnetic north, is assumed to have a random uniform distribution around the circle (\( 0°–360° \)). In other words, the whale is equally likely to be oriented in any horizontal direction anywhere within the specified circle.

To sample depth (\( z \)) and whale pitch (\( \beta \)), it is assumed that Blainville’s beaked whales produce clicks at depths below 300 m during deep foraging dives (Tyack et al., 2006). The vocal period is divided into three phases: descent, foraging, and ascent. The three phases and the proportion of clicks produced in each of them (Table I) are estimated from dive profiles with corresponding vocal periods that are found in the literature (Tyack et al., 2006; Johnson et al., 2006; Ward et al., 2008). It is assumed that during each phase, a whale has certain distributions of depth and pitch. Depths are described by Gaussian distributions with mean and standard deviations estimated from published dive profiles (Johnson et al., 2004; Madsen et al., 2005; Tyack et al., 2006; Ward et al., 2008; Jones et al., 2008). Whale pitch is measured from \( +90° \) to \( -90° \), where negative angles correspond to downward orientation. Hence, during the descent...
phase, the mean pitch is downward (negative); during foraging, closest to horizontal; and during ascent upward (positive). Pitch is described by a beta distribution for the ascending and descending dive phases, with scaling parameters (Table I) that allow for the proportion of clicks to increase for decreasing angles in the descent phase, and for increasing angles in the ascent phase. During the foraging phase, pitch is described by a circular, or von Mises, distribution for which most pitch values are centered on the horizontal, but are also allowed to be oriented up or down (Table I).

The heading of the whale with respect to the hydrophone also matters; it is assumed that the whale can be located in any direction $\phi$ around the hydrophone, where $\phi$ is sampled uniformly in the circle ($0^\circ$–$360^\circ$), with $0^\circ$ being north (Fig. 2). We also define the elevation angle $\eta$ (Fig. 2) between hydrophone and whale as

$$\eta = \arctan\left(\frac{z_b - z}{r_{hw}}\right),$$

where $z_b$ is the depth of the hydrophone, $z$ is the depth of the whale as defined above, and $r_{hw}$ is the horizontal distance between whale and hydrophone.

Animal position and orientation in space with respect to the hydrophone is required to calculate the acoustic off-axis angle $\theta$. This is used to estimate the source level off-axis attenuation due to the animal’s beam pattern and is discussed in Sec. IV C. Following the approach described by Zimmer et al. (2008) and the assumptions made above, the direction vector of the whale is described as

$$a = \begin{pmatrix} a_x \\ a_y \\ a_z \end{pmatrix} = \begin{pmatrix} \sin \gamma \cos \beta \\ \cos \gamma \cos \beta \\ \sin \beta \end{pmatrix},$$

and the direction vector of the hydrophone, as seen by the whale, as

$$b = \begin{pmatrix} b_x \\ b_y \\ b_z \end{pmatrix} = \begin{pmatrix} \cos \eta \sin \phi \\ \cos \eta \cos \phi \\ \sin \eta \end{pmatrix}.\quad (9)$$

The off-axis angle $\theta$ is then calculated by taking the dot product of $a$ with $b$,

$$\cos \theta = a^Tb.\quad (10)$$

Estimations of animal vocalizing depths and distances from the receiving sensor are also used in calculating sound transmission loss (Sec. IV D) between source and receiver, and ultimately the probability of detection with distance.

### B. Source level

The on-axis source level is a critical parameter for estimating the SNR of a click using the sonar equation [Eq. (2)]. However, a literature search for source levels of beaked whales reveals only a few estimates, summarized in Table II. All are derived from either acoustic tag data (Zimmer et al., 2005b), inferences based on tag data (Johnson et al., 2004), or comparisons with echolocation signals from well-studied delphinid species (Madsen et al., 2005). Tyack et al. (2006) suggested that Cuvier’s and Blainville’s beaked whales have similar source levels but different source apertures due to size differences, and that Blainville’s have less directional clicks. It should be noted that source levels are not constant (Zimmer et al., 2008; Madsen et al., 2005), and that the average levels reported also contain uncertainty in their estimations. Another important factor to consider when using source level estimates to calculate the SNR of received levels, which are then compared to measured SNRs, is the time window in which the source level measurement was made. Therefore, a more general parameter to describe source level is source energy flux density [$SL_{E,F}$, used in Eq. (5)]. Zimmer et al. (2005b) measured $SL_{E,F}$ from tagged Cuvier’s beaked whales and reported a maximum level of 164 dB re 1 $\mu$Pa$^2$ s. For modeling purposes a distribution of levels that reflect the uncertainties in measurement is needed. A Gaussian distribution of source energy flux densities was assumed with mean 164 dB re 1 $\mu$Pa$^2$ s and standard deviation 1.769 dB re 1 $\mu$Pa$^2$ s. The standard deviation was estimated by measuring the three highest values from Fig. 6 (left) in Zimmer et al. (2005b), which are believed to be on-axis levels from a single scan of the whale’s sonar beam across the hydrophone.

### C. Off-axis attenuation of source level

The off-axis angle is the angle between the acoustic axis of the whale and the ray that reaches the receiving sensor. For simplicity, the acoustic axis is often assumed to be aligned with the animal’s body axis (Zimmer et al., 2005b). However, the acoustic and body axes may not coincide;

<table>
<thead>
<tr>
<th>Dive Phase</th>
<th>Click Proportion</th>
<th>Mean Depth</th>
<th>Pitch Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Descent</td>
<td>0.20</td>
<td>750 m (+50)</td>
<td>Beta ($-90^\circ \leq \beta \leq 0^\circ$) ($a=2, b=5$)</td>
</tr>
<tr>
<td>Foraging</td>
<td>0.72</td>
<td>1100 m (+50)</td>
<td>Circular ($\beta=0^\circ$) ($\kappa=1$)</td>
</tr>
<tr>
<td>Ascent</td>
<td>0.08</td>
<td>950 m (+50)</td>
<td>Beta ($60^\circ \leq \beta \leq 90^\circ$) ($a=3.5, b=0.9$)</td>
</tr>
</tbody>
</table>

**TABLE I. Values used to estimate whale clicking depth $z$ and pitch $\beta$ for three distinct dive phases. Click proportion is an estimate of the fraction of clicks produced during each dive phase. Mean depth and standard deviation (in parentheses) are used in Gaussian distributions for sampling clicking depths. Pitch distribution indicates the distribution used for each dive phase, the pitch range, and values used for the required distribution parameters. Negative pitch is measured downwards.**

**TABLE II. Source levels of beaked whales, *Ziphius cavirostris* (Zc) and *Mesoplodon densirostris* (Md), available in the literature. Source levels are in dB re 1 $\mu$Pa at 1 m. Note that RMS values can vary depending on the window length used.**

<table>
<thead>
<tr>
<th>Source Level</th>
<th>Species</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>200 dB$_{rms}$</td>
<td>Zc, Md</td>
<td>Zimmer et al. (2008)</td>
</tr>
<tr>
<td>199 dB$_{rms}$</td>
<td>Zc</td>
<td>Tyack et al. (2006)</td>
</tr>
<tr>
<td>Up to 214 dB$_{pP}$</td>
<td>Zc</td>
<td>Zimmer et al. (2005b)</td>
</tr>
<tr>
<td>220 dB$_{pP}$</td>
<td>Md</td>
<td>Madsen et al. (2005)</td>
</tr>
<tr>
<td>200–220 dB$_{pP}$</td>
<td>Zc, Md</td>
<td>Johnson et al. (2004)</td>
</tr>
</tbody>
</table>

**Küsel et al.: Cetacean density estimation—Single sensors**
Zimmer et al. (2005b) point out that both Cuvier’s and Blainville’s beaked whales have been observed to move their heads, and hence the acoustic axis, separately from their bodies while swimming.

Here, the model for a circular piston was used to estimate the attenuation of a beaked whale’s source level as a function of its off-axis angle with respect to the hydrophone. Such a model has been used to estimate the acoustic beam pattern of odontocete species including dolphins (Au, 1993; Rasmussen et al., 2004), sperm whales (Mohl et al., 2003; Zimmer et al., 2005a), and beaked whales (Zimmer et al., 2005b, 2008). For beaked whales, Zimmer et al. (2005b) have shown reasonably good agreement between the piston model and measured apparent source levels from two tagged Cuvier’s beaked whales.

Sound radiated by a circular piston is mathematically described by (Zimmer et al., 2005b; Au and Hastings, 2008)

\[
P(x) = \frac{2P_0 J_1(x)}{x} \quad \text{and} \quad x = \frac{2\pi a \sin(\theta)f}{c},
\]

where \( P \) is pressure, \( P_0 \) is a reference source level, \( J_1 \) is the Bessel function of the first kind and order 1, \( a \) is the piston radius, \( \theta \) is the off-axis angle, \( f \) is frequency, and \( c \) is the speed of sound in seawater. In this analysis, \( P_0 \) was set to 200 dB re 1 \( \mu \)Pa at 1 m (see Table II), \( a \) was assumed to be 16 cm which implies a head diameter of 32 cm, \( \theta \) ranged from \( +90^\circ \) to \(-90^\circ \), \( f \) spanned the 24 to 48 kHz frequency range, and \( c \) was set to 1500 m/s.

The broadband beam pattern for Blainville’s beaked whales is then approximated by integrating Eq. (11) with respect to frequency, using a Gaussian weighting function to account for the variation of source level with frequency (Zimmer et al., 2005b). The weighting function a center frequency of 38.3 kHz and RMS bandwidth of 6.9 kHz was calculated, where incoherent means that the phase of the received signal is random. Both fields were also compared to the common used spherical spreading law, which approximately agrees with the symmetry about 0° (on-axis) can be observed. It is also observed that off-axis clicks will be attenuated by as much as 40 dB. According to Johnson et al. (2006), if Blainville’s have a similar beam pattern to delphinids and Cuvier’s beaked whales, their on-axis clicks will be at least 30 dB stronger than their weakest off-axis clicks.

D. Transmission loss

Sound transmission loss (TL) was calculated using the Gaussian beam tracing model Bellhop (Porter and Bucker, 1987) to model the high frequency clicks of beaked whales. Even though the clicks are broadband in nature, for simplicity a center frequency of 38.3 kHz (Johnson et al., 2006) was used in the transmission loss calculations. A spring speed profile representative of the conditions encountered at AUTEC when the data set was recorded is used as input in the propagation model, and is shown in Fig. 4. The bottom properties for the AUTEC basin, which were also required as input parameters, were adapted from Kennedy and Szlyk (1989). The bottom sound speed, density, and attenuation used in the calculations were 1535 m/s, 1.93 g/cm³, and 0.014 dB/m/kHz, respectively. Use was made of the acoustic reciprocity principle, which states that the received signal will not change when the locations of source and receiver are interchanged in an unchanging environment (Kinsler et al., 1999). Therefore, for ease of calculation, the hydrophone is assumed to be the source and the modeled whale locations are the receivers. Within an 8 km radius of the hydrophone, bathymetric features are smooth. Therefore a single propagation run is performed in a randomly chosen direction from the hydrophone. The resulting matrix of TL values as a function of range and depth is used as input to Eq. (5).

At high frequencies, the sound field is very sensitive to the properties of the medium, and the interference patterns are less stable than at low frequencies (Jensen et al., 1994). Hence, both coherent and incoherent acoustic fields were calculated, where incoherent means that the phase of the pressure, or the interference pattern, is ignored, giving a smoother result. Both fields were also compared to the commonly used spherical spreading law, which approximates
transmission loss by $20 \log(r)$, where $r$ is range expressed in meters. However, at the high frequency modeled, loss due to absorption is a dominant factor (Urick, 1983). This was incorporated into the spherical spreading approximation by adding the frequency-dependent absorption coefficient $\alpha_f$ expressed in dB/km, which gives

$$ TL = 20 \log(r) + \alpha_f (r/1000). \tag{12} $$

A value of 8.9 dB/km (corresponding to a frequency of 38.3 kHz) was used for $\alpha_f$ in the above approximation (Marques et al., 2009). This absorption coefficient was calculated using the Francois-Garrison equation (Francois and Garrison, 1982) integrated from the depth of the whale to the hydrophone using pH, temperature, and salinity measured at AUTEC. Figure 5 shows transmission loss calculated using Eq. (12) together with Bellhop’s coherent and incoherent results. Absorption was also incorporated into the Bellhop model, but through the frequency-dependent Thorp attenuation formula (Thorp, 1967). In Fig. 5 transmission loss curves as a function of range are shown for three different receiver depths corresponding to possible depths where beaked whales produce foraging clicks. Overall, it is observed that there are no significant differences in the transmission loss patterns and magnitudes among coherent, incoherent and spherical spreading calculations. Differences in TL observed at horizontal distances of less than 500 m are not regarded as important as the area surveyed within 500 m is very small compared with the total area surveyed, so inaccuracies here will have small effect on estimates of average detection probability. At ranges greater than 7 km, attenuation is so large that clicks are unlikely to be detected and no significant difference is observed for different depths. These results suggest that for areas where the environment is relatively constant in space and time, a simple spherical spreading law, plus absorption in high frequency cases, should be sufficient for estimating the transmission loss.

**E. Ambient noise levels**

The Tongue of the Ocean represents an acoustically quiet environment due to deep water and the enclosed nature of the site. The main source of ambient noise is wind and wave activity at the surface. It has been observed that ambient noise levels decrease with increasing depth and frequency, most likely due to frequency-dependent acoustic absorption.

Ward et al. (2010) measured ambient noise power between 24 and 48 kHz on a subset of AUTEC’s hydrophones, taking into account hydrophone sensitivity and system gain. These data were found to correlate well with the ambient noise spectrum predicted from historical wind speed measurements. However, at higher frequencies, the measured levels were found to be higher than predicted due to the electronic noise floor of the system. The ambient noise characterization was performed for the 6-day period considered in this study. The data analyzed consisted of manually picked periods of time with minimal biological or man made noise present. Figure 5 shows the 5 different hydrophones, labeled in gray, from which ambient noise measurements were used. They were chosen due to their close proximity to each other and to the single hydrophone used in this work (also shown in Fig. 1). The noise levels used in the analysis from all 5 hydrophones are presented in Fig. 6. It is observed that most of the measured ambient noise levels lie within an almost constant line between 64 and 66 dB re 1 $\mu$Pa. The gaps in the noise levels, where no measurements are observed, are due to the way the data set was divided for analysis and the exclusion of some subsets that presented various problems in the recordings (Marques et al., 2009).

**F. Detector characterization**

To estimate the probability of detecting clicks received at a sensor, it is necessary to characterize the performance of the detector. This was done by measuring the probability of detecting a click as a function of SNR using the 6-day data set recorded on hydrophone 57. Here, detector performance

![FIG. 5. Transmission loss as a function of range, or distance between source and receiver, for 3 different depths: (top) 400, (middle) 800, and (bottom) 1200 m. The different curves represent results of calculations for the coherent (gray) and incoherent (solid black) acoustic fields as well as the spherical spreading law with frequency-dependent attenuation $\alpha = 8.9$ dB/km (dashed).](image1)

![FIG. 6. High-frequency ambient noise levels, integrated over 24–48 kHz, measured at five different hydrophones at AUTEC and used in the Monte Carlo simulations.](image2)
G. Probability of detection: Monte Carlo simulations

The information described above was combined in a Monte Carlo simulation to estimate the average probability of detection as a function of distance. This simulation randomly samples from distributions of the sonar equation parameters described above, then uses these sampled values to calculate the SNR of received clicks using Eqs. (5) and (6). Source level and detector characterization were randomly sampled 1000 times to incorporate the uncertainty of these parameters in the estimated detection probability uncertainty. Each sample of source level and detector performance were incorporated to a set of 10 000 randomly simulated clicks, with coordinates (r, z), and with off-axis angles calculated from the model of animal position and orientation with respect to the hydrophone. For each simulated click, the SNR at the receiving sensor was calculated, and then the detector characterization was used to estimate the probability of detecting the click. At this stage, if the simulated SNR was lower than the minimum observed in the data, the respective click was considered to have 0 probability of detection. At the other extreme, if the simulated SNR was higher than the maximum observed value, the probability of detection for the respective click was assigned the probability of detection of the highest observed SNR. Simulated SNRs between these extremes were assigned the probability of detection corresponding to the detector characterization shown in Fig. 7. The probability of detection was then averaged over all the 1000 × 10 000 simulated clicks to arrive at a value of \( P = 0.0139 \), with SE of 0.0025 and corresponding CV of 17.89%. Plots of probability of detection as a function of slant range for fixed off-axis angles (0°, 45°, and 90°) are shown in Fig. 8 and may be compared to similar plots of Marques et al. (2009, Fig. 3). These results were obtained by
using the coherent TL results from Bellhop in the sonar equation [Eq. (2)].

V. ESTIMATION OF CLICK PRODUCTION RATE

The click production rate \( \hat{r} \), or specifically the mean number of times per second that beaked whales produce clicks, is also an essential variable for this density estimation method [Eq. (1)]. In the case study presented by Marques et al. (2009), Blainville’s beaked whale acoustic tag data from the AUTEC range was used to estimate this quantity. Here, information on animal vocal behavior available in the literature was used to estimate click production rate. It should be noted however, that most of this information is also based on acoustic tags.

The click rate must be averaged over the time period in which density is estimated and should include silent times, that is, times when regular clicks are not observed (Douglas et al., 2005). For beaked whales, periods when no vocalizations are produced include periods at the surface between dives, periods in a foraging dive when clicking pauses during air recycling, periods when buzz clicks are produced (recall that only regular clicks are used as cues), and periods when whales are not making foraging dives. Table III summarizes data gathered from beaked whale publications that was used to estimate the click production rate.

As seen in Table III, several studies have measured inter-click intervals (ICIs). However, such information cannot be inverted into click production rate since it does not account for periods of silence both during deep foraging dives (for example, while emitting buzz clicks) and at or near the surface. Here, click production rate is estimated as

\[
\hat{r} = \frac{\bar{r}_d \times d}{86400},
\]

where \( \bar{r}_d \) is the average number of clicks per whale per deep foraging dive, \( d \) is the average number of deep dives per day, and 86,400 is the number of seconds per day. The former quantity is estimated as

\[
\bar{r}_d = \frac{(\bar{t}_s - \bar{t}_v)}{\text{ICI}},
\]

where \( \bar{t}_v \) is the average duration between times when a whale starts and stops clicking during a deep dive, \( \bar{t}_s \) is the average total time that a whale is either silent or producing buzz clicks during a deep dive, and ICI is the average inter-click interval during the vocal periods. The first quantity, \( \bar{t}_v \), is estimated for Blainville’s beaked whale in three papers (Table III). In cases where a range is given (Johnson et al., 2006), we approximate it by averaging the low and high ends of the range. DiMarzio et al. (2008) give separate estimates for three individuals. Therefore we use seven available estimates for \( \bar{t}_v \): 23, 33, 26, 36, 34.8, 139, and 34.02 min, yielding a mean of 31.17 min with standard error (SE) of 1.83. For average silent period \( \bar{t}_s \), we assume that buzz clicks are the only cause and ignore small silent periods due to air recycling and other reasons. A mean buzz length of 2.9 s is reported by Johnson et al. (2006), and we estimate the number of buzzes per dive to be the average of the three values available in Table III (23, 26, and 38). This yields an estimated \( \bar{t}_s \) of 1.40 min with \( \text{SE} = 0.22 \). For inter-click interval, ICI, six estimates are available from four papers, with a mean of 0.33 s and SE of 0.048. Combining these estimates.

<table>
<thead>
<tr>
<th>Reference</th>
<th>ICI (s)</th>
<th>Click Duration (μs)</th>
<th>Dive Duration (min)</th>
<th>Vocal Time/dive (min)</th>
<th>IDI (min)</th>
<th>Buzzes/dive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frantzis et al. (2002)</td>
<td>~0.4–0.5 (Zc)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Johnson et al. (2004)</td>
<td>~0.4 (Zc)</td>
<td>175 (Zc)</td>
<td>50 (Zc)</td>
<td>–</td>
<td>–</td>
<td>27 (Zc)</td>
</tr>
<tr>
<td>Madsen et al. (2005)</td>
<td>0.2–0.4 (Md)</td>
<td>250 (Md)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>23 (Md)</td>
</tr>
<tr>
<td>Johnson et al. (2006)</td>
<td>0.37 (Md)</td>
<td>–</td>
<td>23–33 (Md)</td>
<td>–</td>
<td>–</td>
<td>26–38 (Md)</td>
</tr>
<tr>
<td>Tyack et al. (2006)</td>
<td>–</td>
<td>–</td>
<td>33 (Zc)</td>
<td>121 (Zc)</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>–</td>
<td>–</td>
<td>26 (Md)</td>
<td>139 (Md)</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Ward et al. (2008)</td>
<td>0.31 (Md)</td>
<td>–</td>
<td>57 (Md)</td>
<td>36 (Md)</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>DiMarzio et al. (2008)</td>
<td>–</td>
<td>–</td>
<td>34.8 (Md)</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>–</td>
<td>–</td>
<td>31.4 (Md)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>–</td>
<td>–</td>
<td>34.02 (Md)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td></td>
</tr>
</tbody>
</table>
in Eq. (14) yields \( \hat{\beta} = 5413.203 \). Standard methods can be used to obtain the variance of a combination of independent quantities (Seber, 1982), and these yield an SE of 858.64 or CV of 15.86%.

The quantity \( \hat{d} \) is estimated as the inverse of the inter-dive interval, for which there is only one estimate (Table III) for Blainville’s beaked whale of 139 min or 0.096 days. This gives an estimate of number of dives per day of 10.36, with no variance. This is in agreement with a mean value of 10 dives per day given by Tyack et al. (2006). Substituting these quantities into Eq. (13) gives an estimate of click production rate \( \hat{r} \) of 0.649 clicks/s, with CV 15.86%. For comparison, the value found by Marques et al. (2009) was 0.407 clicks/s, with CV 9.8%.

VI. RESULTS AND DISCUSSION

Table IV presents a summary of the values for each component of Eq. (1) derived in the previous sections. Combining these results in Eq. (1) then yields an overall population density \( \hat{D} = 69.60 \text{ animals/1000 km}^2 \) (CV 24.64%, 95% CI 50.86–95.25). The estimated average probability of detection \( \hat{P} \) was calculated by using the coherent TL results from the Bellhop propagation model. For comparison, \( \hat{P} \) and the corresponding CV obtained by using the incoherent TL and spherical spreading law were 0.0158 and 18.54%, and the corresponding CV obtained by using the incoherent TL and spherical spreading law were 0.0158 and 18.54%, and the corresponding CV obtained by using the incoherent TL and spherical spreading law were 0.0158 and 18.54%.

The results obtained by this study can be compared to those of Marques et al. (2009). The estimated cue rate \( \hat{r} \), 0.649 clicks/s, is a factor of 1.595 higher than the rate reported by Marques et al. (2009) from direct DTAG data analysis (0.407 clicks/s). The estimate here was based on sparse literature for this species, including little information from this geographic area (most of it was from the Mediterranean Sea). The simulated average probability of detection \( \hat{P} \) was 0.0139, while that estimated analytically by Marques et al. (2009) was 0.032, or 2.30 times higher than the current study. The mean density estimate \( \hat{D} \) of 69.60 animals/1000 km\(^2\), compared to that of Marques et al. (2009) (25.3–22.5 animals/1000 km\(^2\)), differs by a factor of 2.75–3.09. The estimate of Marques et al. (2009) was derived from the whole range, using clicks detected on a total of 82 hydrophones, to obtain values of \( n_i \) and \( \hat{c} \). Re-analyzing the Marques et al. (2009) estimate using the \( n_i \) and \( \hat{c} \) for only hydrophone 57, but values of \( \hat{P} \) (0.032) and \( \hat{r} \) (0.407) from their paper, yields an estimate of 48.164 animals/1000 km\(^2\) (CV 35.08%), differing from the current study’s result by a factor of 1.44.

A detection function for Blainville’s beaked whales was estimated from the Monte Carlo model and is plotted as a function of slant range for fixed off-axis angles (0°, 45°, and 90°) in Fig. 8. The plateau observed in the first 1.5 km for on-axis clicks is due to the assumption that simulated SNRs that are higher than the maximum observed value have probability of detection equal to that of the highest observed SNR. This assumption is an oversimplification and ideally a better characterization of the probability of detection for higher SNRs could be derived from manual analysis of more data. Another option would be to extrapolate the GAM results (Fig. 7), continuing the trend for decreasing detection probability with higher SNR. However, GAM extrapolation is usually unreliable. The results shown in Fig. 8 indicate that off-axis clicks are most likely to be detected within 1 km of the hydrophone, with a rapid decrease in the probability of detection out to 2 km. For on-axis angles, the highest probability of detection occurs at approximately 3 km away from the hydrophone. This peak occurs because of the peak in the curve characterizing detection probability as a function of SNR (Fig. 7). The probability of detection then decreases rapidly with increasing distance, reaching almost zero probability at approximately 5 km. The estimated detection functions from Marques et al. (2009) (Fig. 3, upper right corner) show overall higher probabilities of detection at shorter slant distances as compared to this study. It should be noted that in this study the off-axis angle was calculated by implementing a piston model to approximate the beam pattern of Blainville’s beaked whale, and such an approach does not discriminate the components of the off-axis angle into vertical and horizontal.

Results of our calculations show that the higher the energy flux density (or source level) used (from the defined normal distribution), the higher the average probability of detection for the 10 000 simulated clicks (Fig. 9). This

![FIG. 9. Average probability of detection from the 10 000 simulated clicks for each realization of energy flux density taken from a Gaussian distribution.](image-url)
suggests that uncertainty in source levels contributes heavily
to variance in calculated density, as compared to uncertainty
due to detector characterization. At the same time, Fig. 7
also shows considerable uncertainty about the probability of
detection at high SNR values. It is stressed here that the de-
tector-and-classifier characterization performed in this study
relied on only a single, small, randomly chosen time period.
Detection characterization for future density estimates
should be more systematic and extensive. We also note that
the use of this density estimation method requires a detector
that is well characterized rather than optimal. Uncertainty
associated with the detector performance was measured by
the residual variance of the generalized additive model
(GAM) fit to the detector characterization data. This uncer-
tainty is then used in the final density estimate uncertainty
by the Monte Carlo procedure: each iteration of the simul-
ation used a random detector characterization curve, sampled
from a list of possible curves derived from the GAM. So
uncertainty in density is not dependent on how good or bad
the detector is, it depends on how well the detector is charac-
terized, and hence what the residual error on the GAM curve
is. However, a poor detector will give a lower number of
detections to work with, and this will contribute to a higher
variance in density (as the encounter rate CV will be higher).
Hence, even if it can be accurately characterized, having a
poor detector is not without cost. To obtain accurate density
estimates in future acoustic monitoring studies, accurate esti-
mation of cue (click) rates is also required. For sperm
whales, for example, it is known that group size (Whitehead
and Weilgart, 1990) and time of day (Weilgart, 1991) affect
whales, for example, it is known that group size (Whitehead
terations and animal vocal behavior often vary with geographic
location and target species. Previous density estimates of
Blainville’s beaked whales at AUTEC have relied on differ-
tent analyses to arrive at estimated densities. Moretti et al.
(2006) and Marques et al. (2009) both relied on DTAG data
collected at AUTEC simultaneously with their sound record-
ings to arrive at their density estimates. Here only informa-
tion found in the literature was used to estimate density.
Note that the information found in the literature for beaked
whales comes mostly from DTAG data, although none of the
data used are from the geographic area of this study. By only
using information available in the literature we hope to point
the way toward density estimates that may be made with
minimal, or even no use of acoustic tag data at all. The idea
is that such data are sufficiently difficult and expensive to

collect that they cannot be used for every species, and in ev-
every area, where density estimates are desired.

When counting cues or estimating the rate of false posi-
tive detections, multipath arrivals can be a source of error
and must be accounted for during data analysis. In the case
of beaked whale clicks at AUTEC, recordings of vocalizations
rarely have detectable multipath arrivals, perhaps due
to the highly directional nature of the clicks. However, it is
possible to separate direct and indirect arrivals using the dif-
f erent signal characteristics of the two arrivals (Baggenstoss,
2008). For sperm whales and dolphins, which more com-
monly have multipath arrivals, grouping clicks into trains
(Starkhammar et al., 2011; Baggenstoss, 2011) can also help
in identifying direct and reflected-path arrivals. Another al-
ternative is to use manual validation by a trained analyst to
determine the average number of times clicks are received,
and then incorporate this information in the estimation of
the false positive rate. Although difficult, this is feasible under
certain circumstances, as, for example, when few animals
are present or when vocalizations are relatively uncommon.

Estimates of variance and confidence limits are just as
important as the estimate of the mean. Our estimates of
uncertainty for several relevant parameters were based on
poorly supported assumptions about real-world uncertainty
in the input quantities. For example, for click rate, we used the
variance from an unweighted average of reported estimates,
which probably underestimates true uncertainty. For source
level and animal beam pattern, rough approximations were
made from assumptions of those parameters found in the lit-
erature. Here a single sensor was considered, a conscious de-
cision given the proof-of-concept goal of this work. But the
use of a single sensor implied that the variance in the number
of cue counts had to be obtained based on a distributional
assumption. While a Poisson distribution was assumed, in
other similar settings an over-dispersed Poisson has been sug-
gested (e.g., Buckland et al., 2001), and so there is not a defi-
nite answer regarding the variance estimation under this
setting. However, more robust estimates would be based on a
design that considers an array of hydrophones even if, for
practical purposes, they are all operating independently. This
will allow variance on counts over space to be obtained using
an empirical variance estimator. Note that this is the recom-
ended procedure in closely related distance sampling sur-
veys (Buckland et al., 2001), and the one we recommend to
obtain reliable estimates of density. Additionally, an array of
sensors means that, averaged across space, the assumption
of uniform distribution in space, used here to obtain the esti-
mated average probability of detection, is much more likely
to hold.

Population density estimation based on passive acous-
tics is in its infancy. However, the approach looks promising.
We anticipate considerable research in this area: developing
dedicated hardware, refining detection and classification
algorithms, further developing estimation methods, and
improving our knowledge of species’ acoustic behavior. As
an example, cue rate estimates are needed in general to con-
vert detected cues to a measure of abundance, but such rates
are currently unknown for most species. In general, directly
measuring the variables required for density estimation is far
preferable to deriving their values based on assumptions and modeling. For example, estimation of a detection function based on in situ data measurements (e.g., Marques et al., 2009) or dense arrays of linked sensors (e.g., Marques et al., 2010) automatically integrates all species/site specific characteristics, and avoids the need for assumptions made in this paper about source levels, directionality, and sound propagation. However, in situ measurements and dense linked sensor arrays can be prohibitively expensive. Our results show that it is possible to use a model-based approach to produce density estimates in reasonable agreement with those from other methods, by making careful use of information available in the literature.

Extensive passive acoustic recordings exist of marine mammal vocalizations from single fixed sensors or sonobuoys. The method described in this paper opens up the opportunity for estimating animal density by analyzing such previously recorded datasets, contributing to the understanding of cetacean population distribution and behavior. We hope that such methods are found to be useful in the common situation where unlinked sensors are deployed, and information from tagged animals is not available for the time and place where acoustic monitoring took place.

ACKNOWLEDGMENTS

This research was undertaken as part of the Density Estimation for Cetaceans from passive Acoustic Fixed sensors (DECAF) project, work supported by National Oceanographic Partnership Program grant 2007-0145-002. It was also supported by Naval Postgraduate School grants N00244-08-1-0029, N00244-09-1-0079, and N00244-10-1-0047. This is NOAA/PMEL contribution No. 3648.


Küsel et al.: Cetacean density estimation—Single sensors

3621


Downloaded 04 Jul 2011 to 138.251.202.200. Redistribution subject to ASA license or copyright; see http://asadl.org/journals/doc/ASALIB-home/info/terms.jsp