Estimating cetacean population density using fixed passive acoustic sensors: An example with Blainville's beaked whales

Tiago A. Marques[a] and Len Thomas
Centre for Research into Ecological and Environmental Modelling, The Observatory,
University of St Andrews, St Andrews KY16 9LZ, Scotland

Jessica Ward and Nancy DiMarzio
Naval Undersea Warfare Center Division, 1176 Howell Street, Newport, Rhode Island 02841

Peter L. Tyack
Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543

Methods are developed for estimating the size/density of cetacean populations using data from a set of fixed passive acoustic sensors. The methods convert the number of detected acoustic cues into animal density by accounting for (i) the probability of detecting cues, (ii) the rate at which animals produce cues, and (iii) the proportion of false positive detections. Additional information is often required for estimation of these quantities, for example, from an acoustic tag applied to a sample of animals. Methods are illustrated with a case study: estimation of Blainville's beaked whale density over a 6 day period in spring 2005, using an 82 hydrophone wide-baseline array located in the Tongue of the Ocean, Bahamas. To estimate the required quantities, additional data are used from digital acoustic tags, attached to five whales over 21 deep dives, where cues recorded on some of the dives are associated with those received on the fixed hydrophones. Estimated density was 25.3 or 22.5 animals/1000 km^2, depending on assumptions about false positive detections, with 95% confidence intervals 17.3–36.9 and 15.4–32.9. These methods are potentially applicable to a wide variety of marine and terrestrial species that are hard to survey using conventional visual methods. © *2009 Acoustical Society of America.* [DOI: 10.1121/1.3089590]

I. INTRODUCTION

Cetaceans (whales and dolphins) form a key part of marine ecosystems, and yet many species are potentially threatened with extinction by human activities. One essential element of an effective conservation or management strategy is a reliable estimate of population size ("abundance") or, equivalently, number per unit area ("density"). However, most cetacean species are hard to survey, since they live at low density over large areas of ocean and spend almost all of their time underwater. The object of this paper is to increase the repertoire of tools available for making species assessments, by developing and demonstrating methods for estimating cetacean density from surveys of their vocalizations collected from fixed passive acoustic sensors.

Currently, the main method for obtaining estimates of density is through visual line transect surveys. A set of randomly placed lines is traversed by an observation platform (e.g., ship, airplane, or helicopter) and all sighted animals of the target species are recorded, together with their perpendicular distance from the line. In the standard method, it is assumed that all animals on the transect line (i.e., at zero distance) are seen with certainty, but that probability of detection declines with increasing distance from the line. The distribution of observed detection distances is then used to estimate the average probability of detection, and this in turn allows estimation of population abundance or density. Line transects are a special case of distance sampling methods, which are described in detail in the two standard texts by Buckland *et al.* (2001, 2004).

Visual line transect methods have a number of disadvantages for surveying cetaceans: they can only be performed during daylight hours and are strongly dependent on good weather conditions; they do not work well for species that spend long periods of time underwater; they are expensive to do well and have restricted temporal coverage. On the other hand, some cetacean species make frequent and characteristic vocalizations, and this has led to increasing recent interest in the use of passive acoustic methods for monitoring cetacean populations (see review by Mellinger *et al.*, 2007b). One solution is to replace or supplement the visual observers on a shipboard line transect survey with a towed passive acoustic platform, since even a simple two-element hydrophone array can be used to obtain locations of repeatedly vocalizing animals, and hence the required perpendicular distances. This has proved particularly effective for sperm whales (*Physeter macrocephalus*), which are long, deep divers and hence hard to detect visually, but produce loud

---

[a]Electronic mail: tiago@mcs.st-and.ac.uk; URL: http://www.creem.st-and.ac.uk/decaf/

echolocation clicks (Hastie *et al.*, 2003; Barlow and Taylor, 2005; Lewis *et al.*, 2007).

Acoustic line transect methods still require an expensive survey platform and have restricted temporal coverage. By contrast, technology exists where fixed acoustic sensors (either anchored or buoyed) can be deployed to record data over long periods or transmit it to shore via cables. Mellinger *et al.* (2007b) provide a review of the potential modalities. Such sensors have the potential to provide relatively inexpensive long-term monitoring and, if deployed in a wide-baseline array, could also cover large spatial areas.

There are several examples of the use of fixed acoustic sensors to provide an index of abundance, or at least to detect the presence of a species over time (e.g., Au *et al.*, 2000; Charif *et al.*, 2001; Moore *et al.*, 2006; Mellinger *et al.*, 2007a; Stafford *et al.*, 2007). However estimating absolute density or abundance from such data is considerably more difficult, although considerably more useful too. Broadly, the key problems are as follows: (i) to convert the number of vocalizations detected to the number produced within a defined area, by accounting for false positive and false negative detections, and (ii) to convert this estimate of vocalizations produced to the average number of animals present. To date, no method has been proposed that might be used in general to estimate cetacean density from fixed passive acoustic devices (Mellinger *et al.*, 2007b), although several papers have addressed elements of the problem (e.g., Aubauer *et al.*, 2000; McDonald and Fox, 1999; Wiggins *et al.*, 2004; Stafford *et al.*, 2007; Zimmer *et al.*, 2008).

This paper presents a framework for estimating cetacean density from an array of fixed passive acoustic detectors. The framework lays its foundation in a type of distance sampling survey methodology called cue counting (Buckland *et al.*, 2001) and is general enough that it might be used under considerably different scenarios, with appropriate modifications that are also discussed. The proposed approach is illustrated by a case study: estimating the density of Blainville's beaked whale, *Mesoplodon densirostris* (Md), at the Atlantic Undersea Test and Evaluation Center (AUTEC) range, an instrumented US Navy testing range in the Bahamas. Part of the data used here were also used in a previous attempt to estimate Md density at AUTEC by Moretti *et al.* (2006), which provides a useful comparison for the results obtained.

II. STATISTICAL FRAMEWORK FOR ESTIMATING DENSITY

Density D is defined as a number of animals per unit area, and hence density can be expressed as

$$D = N/a, \tag{1}$$

where N is the number of animals present in a region of area a. It is therefore straightforward to convert any density estimate into an abundance estimate, as long as the area it applies to is known, as $N = Da$. Given the number of animals n detected in survey area a, an intuitive estimator of abundance is given by

$$\hat{D} = \frac{n}{a\hat{P}}, \tag{2}$$

where $\hat{P}$ represents the estimated probability of detecting an animal. P can be estimated by a number of methods, of which distance sampling (Buckland *et al.*, 2001) is arguably the most commonly used.

A. Proposed estimator and variance

Here it is assumed that a set of replicate fixed passive acoustic sensors has been deployed using a random design (e.g., systematic random) within the survey region of interest (see, for example, Buckland *et al.*, 2001, Chap. 7 and Buckland *et al.*, 2004, Chap. 7, for discussions of appropriate sampling designs). The goal is to estimate average density and/or abundance of a cetacean species over the time period that the sensors were deployed. The methods generalize easily to multiple species, spatial and temporal stratification, etc. They are also applicable to single sensor situations and nonrandom sensor placement, although in these cases caveats must be introduced.

The methods proposed here involve counting (for a known period of time) the number of detected acoustic cues produced by the animals of interest, and appropriately scaling up this number of detected cues to estimate animal density. The appropriate definition of a cue is study-specific, but it will usually be a distinct animal vocalization such as a song unit, echolocation click, etc. Considering Eq. (2), an estimate of density $\hat{D}$ from the n_c detected cues over a time period T (and additional information detailed below) can be obtained by

$$\hat{D} = \frac{n_c(1 - \hat{c})}{K\pi w^2 \hat{P} T \hat{r}}, \tag{3}$$

where w is the distance away from the hydrophones beyond which cues are assumed to not be detected, $\hat{P}$ is the estimated average probability of detecting a cue made within distance w, $\hat{r}$ is the estimated cue production rate, $\hat{c}$ is the estimated proportion of false positive detections, and K is the number of replicate sensors used.

Note that Eq. (3) can be seen as a special case of Eq. (2). $n_c(1 - \hat{c})$ corresponds to the number of detected cues that were actually from the target species. $T\hat{r}$ corresponds to the number of cues produced by an average animal during the recording time T. Hence $n_c(1 - \hat{c})/T\hat{r}$ corresponds to n in Eq. (2), while $K\pi w^2$ corresponds to a in Eq. (2).

Strictly speaking, the hydrophones should be independent sampling units, but in practice some clicks are detected at more than one hydrophone. However, Buckland (2006) showed by simulation that these methods are insensitive to this sort of nonindependence.

Assuming independence of the various random components, the variance for the density estimator can be approximated via the delta method (Seber, 1982) (see also Powell, 2007, for an applied review of the method) as

$$\text{var}(\hat{D}) \approx D^2 \{ CV(n_c)^2 + CV(\hat{r})^2 + CV(\hat{c})^2 + CV(\hat{P})^2 \}, \quad (4)$$

where CV denotes the coefficient of variation of the estimate, i.e., the standard error of the estimate divided by the estimate.

The variance in n_c can be estimated from the empirical variance of the cue counts n_{cj} $(j=1,2,\ldots,K)$ over K hydrophones, as is usually done for conventional distance sampling methods (Buckland *et al.*, 2001, p. 78) as

$$\text{var}(n_c) = K \sum_{j=1}^{K} \frac{[n_{cj} - n_c/K]^2}{K-1}.$$

For the other three components, obtaining the appropriate variance will depend on how these are estimated. Examples are given in the case study.

Confidence intervals on density can be obtained by assuming that the density estimate follows a log-normal distribution (Buckland *et al.*, 2001, p. 77). An alternative approach for obtaining both variance and confidence intervals is the nonparametric bootstrap [Buckland *et al.*, 2001; see section 3.6 for details in this context, and Efron and Tibshirani, 1993 for an overview]. Both the analytic and bootstrap methods are commonly employed in conventional distance sampling.

Sections II B–II D consider methods for obtaining estimates of the three components P (probability of detection), r (cue production rate), and c (false positive proportion).

B. Probability of detection

To understand the potential methods for estimating P, it is convenient to start by describing the estimation of the detection function, $g(y)$, that is the probability of detecting a cue given it is produced at horizontal distance y. Here y refers to horizontal distance (i.e., projected onto the sea surface); the use of direct (i.e., slant) distances is discussed later.

In conventional distance sampling methods, including cue counting, $g(y)$ is estimated using the horizontal distances of detected objects from the line or point (usually referred to as "transect," with the equivalent here being a point with the acoustic sensor at the center, from which acoustic cues are detected). This is feasible because, with random transect placement, the *true* distribution of object distances is known; hence any departure in the distribution of *detected* distances from this known distribution can be attributed to the detection process. Specifically, in variants of distance sampling where the transect is a point, as in the current case, the true distribution of horizontal object distances, $h(y)$, is given by $2y/w^2$, where w is some distance beyond which detections are not recorded (and can be infinity). Thus, the probability density function (pdf) of observed distances, $f(y)$, is given by

$$f(y) = \frac{h(y)g(y)}{\int_0^w h(y)g(y)dy} = \frac{2yg(y)}{w^2 P}, \quad (5)$$

where $P = \int_0^w h(y)g(y)dy$ acts as a normalizing constant that ensures $f(y)$ is a pdf (i.e., that integrates to 1). Note that $g(y)$ can be multiplied by any constant without changing $f(y)$; in other words, the observed distances give information about the shape of $g(y)$ (i.e., the change in detectability with dis-

tance) but not the absolute level of $g(y)$. Therefore, an additional assumption must be made before $g(y)$ can be estimated, and the conventional assumption is that all objects at zero distance are detected, i.e., that $g(0)=1$. Then, given a parametric model for $g(y)$, the parameters of the model can be estimated from a sample of distances using, for example, maximum likelihood estimation.

Once an estimate of the detection function parameters is obtained, it is straightforward to estimate the average probability of detecting an object within w, as the average of the estimated detection probability at each distance multiplied by the density of objects at that distance:

$$\hat{P} = \int_0^w \hat{g}(y)h(y)dy. \quad (6)$$

Specialized software, DISTANCE (Thomas *et al.*, 2006), is freely available for the analysis of such data, and variance estimates and confidence limits are readily available from the output.

The conventional approach is sometimes possible using data from fixed acoustic sensors, so long as (i) horizontal distances to detected cues can be obtained, and (ii) the basic assumptions are met: that the true distribution of horizontal object distances, $h(y)$, is known and that $g(0)=1$. Calculation of horizontal distances using a single bottom-mounted sensor was demonstrated by McDonald and Fox (1999), using transmission loss and multi-path methods to estimate distance to calling fin whales near a Hawaiian seamount, and by Wiggins *et al.* (2004), using waveguide propagation to estimate distances to North Pacific right whales in the Bering Sea. However, the assumption of known $h(y)$ is often questionable with a single sensor. If cues are received on multiple sensors and can be associated, then time differences of arrival can potentially be used to localize the source, thereby providing a distance (e.g., Freitag and Tyack, 1993). However, the conventional methods would then need to be extended to allow the input data to be a set of detections of the same cue from multiple sensors at different distances.

An alternative way to obtain the detection function is to derive it from assumed models for sound production, propagation, and detection. This approach was demonstrated (using different modeling approaches and on different species) by Zimmer *et al.* (2008) and Stafford *et al.* (2007). Such approaches are naturally heavily dependent on the accuracy of the models and input parameters.

A third approach is to estimate the detection function empirically, using a set of experimental "trials" where vocalizations are produced at known times and distances from sensors, and the outcome is whether they are detected or not. One way to achieve this is to play recorded vocalizations, but a more realistic method involves placing acoustic recording tags on a sample of animals that are within the vicinity of the sensors, and using the tag data to determine where and when vocalizations were produced. Ideally, this additional experiment should be undertaken concurrently with the primary survey and on a representative sample of animals and sensors; otherwise differences in conditions between the tagging experiment and the primary survey will need to be accounted

1984   J. Acoust. Soc. Am., Vol. 125, No. 4, April 2009

Marques *et al.*: Cetacean density estimation from passive acoustics

for, by modeling covariates thought to influence detectability and then predicting detectability at the time of the trials. This is the approach taken in the case study, and the analysis options are described more fully later.

C. Cue production rate

The cue production rate r can be estimated using acoustic tag data, as described at the end of the Sec. II B, and this is the approach taken in the case study. Again, ideally, this would be collected concurrently with the primary survey and on a representative sample of animals. Otherwise, factors that potentially affect cue rate such as season, time of day, sex, etc., will need to be accounted for.

A second approach, if additional experimental tagging data are unavailable, is to use a cue rate estimate published elsewhere. For some species, which have very predictable vocalization patterns, this may produce a reliable estimate; for others it will not.

A third alternative, if it is possible to associate successive vocalizations from individual animals within the sensor array (such as when calls are uniquely identifiable or density is very low), and if it can be determined when animals are so close to a sensor that no calls will be missed, is to use the cue rate from vocalizations produced by these animals. However, for species that can pass by a sensor without producing any cues, this will inevitably lead to an overestimate of cue rate.

D. False positive proportion

To be recorded as a cue, a sound must go through a two-stage process: it must be (i) detected as being potentially relevant and (ii) classified as the correct type of sound. Because of the volume of data involved, both stages are typically performed by computer algorithms.

If a sound of interest, i.e., a vocalization by an animal from the target species within distance w of a sensor, is not detected, or is detected but incorrectly classified, it is referred to as a "false negative." (In statistics, this is also known as a type II error, considering the combined detection and classification process to be a statistical test.) The detection function, via the estimated detection probability, corrects for this type of error.

On the other hand, it is possible for a sound that is not a vocalization from the target species to be detected and classified as a cue, creating a "false positive" (or a type I error). The proportion of sounds classified as cues that do not come from the target species is the false positive proportion, and this must be estimated. As with the previous components, there are multiple ways to achieve this.

The most direct approach is to take a sample of the sounds classified as cues and manually check what proportion are false positives. This is the approach taken in the case study, and relies on the human operator to make correct classifications. If a human operator cannot be considered reliable, an alternative is to use some performance statistics of the detector/classifier, for example, on data where animals were known to be vocalizing because they were tagged or being observed from a nearby boat, and predict the average false positive proportion under the conditions of the survey

being analyzed. This requires knowledge of all the variables that affect the false positive proportion, and their distribution both for the dataset where the performance characteristics were tested and for the dataset for which density is being estimated. This seems hard to implement as it implicitly requires one to know the total number of positive and negative sounds (i.e., whale clicks versus sounds from other sources). Hence the first approach is to be preferred wherever it is feasible.

Note in general, one cannot generalize to other scenarios the proportion of false positives in a given place and time. In particular, if the relative abundance of species contributing to false positives changes, the false positive proportion will change accordingly.

III. CASE STUDY: BLAINVILLE'S BEAKED WHALES AT AUTEC

The proposed methods are illustrated here with a detailed case study: the estimation of average density of Blainville's beaked whales, Md, at the AUTEC underwater tracking range during a 6 day period in spring 2005. The inference made is restricted to the time periods, during these 6 days, for which recordings were available (see below for details).

Beaked whales are considered difficult to survey using visual methods, since they make long, deep dives and are relatively difficult to sight even for the short periods of time they are at the surface (Barlow *et al.*, 2006; Barlow and Gisiner, 2006). However, they make regular high-frequency echolocation clicks during the deeper part of their dives (Johnson *et al.*, 2004) which, although highly directional (Zimmer *et al.*, 2005), can be detected by both surface-and bottom-mounted hydrophones, with the latter likely to have better detection ranges (Zimmer *et al.*, 2008). Md have previously been detected up to 6500 m slant range from the AUTEC hydrophones (Ward *et al.*, 2008).

AUTEC is an instrumented US Navy testing range located in the Tongue of the Ocean, Bahamas. It contains a wide-baseline array of 93 cabled hydrophones, mounted 4–5 m off the sea floor at depths between 1.3 and 2 km, and with an upward, roughly hemispherical beam pattern. For the data presented in this study, 82 hydrophones were digitally recorded at 96 kHz (Fig. 1): 68 wideband hydrophones spaced 4 km apart with a usable bandwidth from 50 Hz to 45 kHz, and 14 hydrophones spaced 1.8 km apart with a bandwidth from 8 to 50 kHz (bandwidth and spacing values are approximate). Previous research utilizing this hardware has demonstrated the presence of Md at AUTEC, developed suitable automated detectors and estimated approximate animal density (Moretti *et al.*, 2006; DiMarzio *et al.*, 2008; Ward *et al.*, 2008).

The processed data and metadata used in this paper were submitted to the Ocean Biogeographic Information System and are accessible via http://seamap.env.duke.edu/.

A. Case study data and data processing

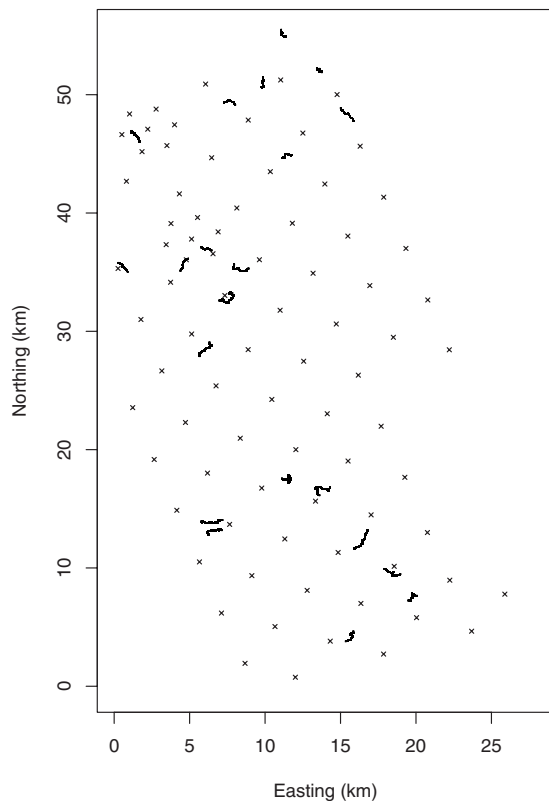Two sets of data were used in the case study. The first (the "primary dataset") was a set of recordings taken on 82

FIG. 1. The spatial layout of the AUTEC hydrophones that were recording during the collection of the primary survey data, represented by small crosses. Also shown as dots (perceived as solid lines) are the locations of the tagged Blainville's beaked whales when each click recorded on the DTag was produced.

TABLE I. Details about the tagged whales used in case study analysis: Tag ID, date the animal was tagged, number of dives while animal was tagged, and number of dives with data available for estimating the detection function $g(y)$.

Tag	Date	Number of Dives	Dives for $g(y)$
Md296	23 Oct 2006	3	3
Md227	15 Aug 2007	6	0
Md245	2 Sep 2007	4	3
Md248a	5 Sep 2007	4	4
Md248b	5 Sep 2007	4	3
Total		21	13

cessed to estimate the location of the whales at each click event, and these clicks were matched with clicks received on surrounding AUTEC hydrophones to enable estimation of the detection function.

A brief description follows of the processing required to extract detections from the main dataset, and to extract click rates, locations, and associations with AUTEC hydrophones from the auxiliary dataset.

1. Processing of primary dataset

Clicks were detected using a multi-stage fast Fourier transform (FFT)-based energy detector (Morrissey *et al.*, 2006), using data digitally recorded from the range hydrophones at a 96 kHz sampling rate. A 2048-point FFT with 50% overlap was used. This provides a 46.875 Hz frequency resolution and a 10.67 ms time resolution. During the first stage, an adaptive threshold was run in each bin of the FFT. If energy above threshold was detected, the bin was set to a "1" and a detection report was generated. During the second stage, detections are classified based on the frequency distribution of the detected signal. Detections are classified as beaked whale clicks by evaluating the ratio of bins above threshold to the total number of bins within the frequency band of 24–48 kHz (Morrissey *et al.*, 2006; Ward *et al.*, 2008). Additionally, a second check is performed by examining the number of bins set outside the 24–48 kHz "beaked whale band." If this exceeds 10%, then the detection is reclassified as a dolphin as they are more likely to have significant spectral energy below 24 kHz.

The detection reports classified as beaked whales were then summed into a final output consisting of the number of beaked whale clicks per minute for each hydrophone over the recording period.

2. Processing of auxiliary dataset

Additionally, DTag (Johnson and Tyack, 2003) data, both positional and acoustic, were available for five whales, in a total of 21 (Table I) deep dives. The tag records continuously from a built-in hydrophone and suite of sensors that convey information on positional data: animal depth, heading, pitch, and roll. The sensors sample the orientation of the animal in three dimensions with sufficient resolution to capture individual fluke strokes (Johnson and Tyack, 2003). The speed from the pressure sensor, heading, and pitch angle were used to reconstruct a three-dimensional "pseudotrack"

hydrophones during 6 days in spring 2005 (between 17:03 GMT, 26 April and 08:32 GMT, 2 May 2005). These were processed to extract detections of Md foraging clicks (the cues), and the number of detections was used as the basis for estimating average density of Md over the 6 day period. For operational reasons the data were divided into 12 sets, and entire sets (2, 5, 7, and 11), as well as some scattered minutes, were not used. These periods were removed from the analysis for various reasons (e.g., problems with acoustic recordings, system tests being conducted over the range, etc.).

Traditional hyperbolic localization techniques used to obtain the position of sound sources such as whales require a click to be detected on at least four hydrophones. However, previous studies have observed that the combination of Md narrow beam width and AUTEC's wide hydrophone baselines result in clicks often not being detected on more than one or two hydrophones simultaneously (Ward *et al.* 2008); hence detection function and cue rate estimation methods based on localizing Md at AUTEC were considered infeasible. Therefore, a second dataset was required to obtain these quantities. This (the "auxiliary dataset") comes from acoustic tagging studies carried out in October 2006 and August–September 2007 in which five whales were fitted with digital acoustic tag (DTag) acoustic recording tags (Johnson and Tyack, 2003) which remained attached for a total of 21 deep dives (Table I). The tag data were used to estimate the cue rate. A subset of the data was further pro-

for each animal during the period the tag was attached (Johnson and Tyack, 2003). These data were used for the estimation of click rates, and also for simulation of click characteristics involved in the estimation of the mean detection probability (see below).

For the estimation of the detection function, further data processing was required. For four whales and a total of 13 of the deep dives, these DTag data were associated with matched filter detections from the surrounding hydrophones, localized, and the pseudotrack georeferenced (i.e., absolute rather than relative coordinates obtained; see Ward *et al.*, 2008 for further details). Once georeferenced, the DTag accelerometer and magnetometer data for each click within the georeferenced portion of the dive were used to calculate the spatial relationship between the whale and each hydrophone in the whale orientation frame (see Johnson and Tyack, 2003; Ward *et al.*, 2008, for further details). An indicator of detection (1—detected; 0—not detected) could then be generated for each click produced at each surrounding hydrophone (using the same FFT-based detector as in the primary survey click count data), as well as a set of corresponding click characteristics in the frame of the whale (namely, slant distance, vertical, and horizontal off-axis angles with respect to the whale's orientation).

B. Case study density estimation methods

1. Estimating probability of detection

As stated earlier, the conventional distance sampling approach of estimating average detection probability using detection distances could not be used due to the difficulty in easily localizing Md clicks using the wide-baseline hydrophone array at AUTEC. Hence auxiliary data from the tagging study were used, where for each click produced by a tagged whale, it was known which of the surrounding hydrophones detected the click.

If tags had been applied to a large random sample of animals from the population present at AUTEC during the 6 day period of interest, then the average proportion of clicks that were detected would be an unbiased estimator of the average detection probability required. However, the tags were applied in a different time period, and the sample size of tags was not large. For both of these reasons, a more complex approach was required, where the probability of detection was modeled as a function of covariates thought to influence detectability, and the fitted relationship then used to estimate average detection probability during the 6 day period.

The modeling approach used was a generalized additive model (GAM, Wood, 2006), with a binomial response and logistic link function, to model the probability of detecting a Md click as a function of available covariates. The dependent variable used was the outcome of the detection and classification process (i.e., detected and classified as a cue/not detected or not classified as a cue) for each click produced by the tagged whale at each hydrophone within 8 km of the location of the whale (at the moment of the first click in the corresponding dive). The explanatory variables considered were the slant whale-hydrophone distance and the horizontal

(*hoa*) and vertical (*voa*) off-axis angles, measured on the frame of the whale, with respect to the hydrophone (e.g., a 0,0 angle pair corresponds to a whale pointing straight at the hydrophone). For slant distance, a cubic regression spline basis was used, while for the off-axis angles a two-dimensional tensor product based smooth was used, to incorporate interactions between the horizontal and vertical angles. The largest detection distance for a click was at around 6.5 km, and the model was fitted truncating all data for which distance was larger than 8 km. The maximum degrees of freedom for the smooth of slant distance were set to 4 as, *a priori*, a relatively smooth decreasing function of distance was expected. The models were implemented using the library mgcv in R version 2.8.0 (R Development Core Team, 2008).

Given the fitted detection function model, there are, in general, two approaches for using this model to predict average probability of detection for the primary dataset. If the values of the covariates used in the model are also known for each detection in the primary dataset, then detection probability can be predicted conditional on these values and an average taken. However, if the covariate values for each detected click in the primary dataset are not known, but the *distribution* of the covariate values is known, then its effect can be integrated out of the model, giving the correct average detection probability. This is the approach taken in conventional distance sampling with the distance covariate [Eq. (6)].

In the current study, the values of slant distance and orientation were not known for any of the clicks in the primary dataset. Hence, a simulation approach was used to approximate the distribution of these covariates, based on the motion data from the DTags in the auxiliary dataset. The underlying assumptions were as follows: (i) the paths recorded in the 21 DTag dives from the auxiliary dataset are representative of the paths during the primary survey period, and (ii) the initial position of dives is random with respect to hydrophone position during the primary survey period. To implement the simulation, the following procedure was repeated many times:

(1) Generate the hypothetical location of the first click in a deep dive (x_i, y_i); locations were generated randomly in a buffer area of 8 km around AUTEC.
(2) Sample a dive from the 21 DTag dives available and select the data corresponding to the position of the whale in space (x, y, z location, pitch, heading and roll) at each click event.
(3) Relocate that dive such that the first click is at (x_i, y_i).
(4) For each click event, obtain the slant distance, *hoa* and *voa* with respect to all hydrophones closer than a distance of 8 km to the whale position.

In practice, instead of sampling each click event, only every 100th click was sampled as this was found to increase the speed of convergence of the algorithm to the true distribution of the covariate values. With this refinement, the distributions stabilized after 1000 simulated dives.

Given the above samples from the distribution of cova-

riates, the detection function model fitted earlier was used to estimate detection probability for each click, and the mean of this was taken as an estimate of average detection probability, $\hat{P}$.

To estimate the variance in $\hat{P}$ a nonparametric bootstrap procedure was implemented (e.g., Efron and Tibshirani, 1993). In each of 250 bootstrap replicates, a sample of dives was taken with replacement from the set of 13 dives from which the original detection function had been estimated, the GAM was refitted to this bootstrap resample, and the new fitted model used to produce a new estimate of $\hat{P}$. The empirical variance obtained over these bootstrap estimates is an estimate of the original variance of $\hat{P}$ (see pp. 82–83 of Buckland *et al.*, 2001, for further details in a similar context). This approach assumes that the dives are statistically independent.

2. Estimating cue production rate

The cue production rate (mean number of clicks per whale per second) was estimated using the DTag data. To avoid bias, the periods used for estimating the cue rate were restricted to complete deep-dive cycles (from the start of a deep dive to the start of the next), for which both sound and positional data were available, in a total of 21 deep dives from five whales.

The mean click rate was estimated as a weighted mean of click rate per deep-dive cycle, weighted by deep-dive cycle time length. The variance of this weighted mean was estimated using Cochran's approximation as recommended by Gatz and Smith (1995).

3. Estimating false positive proportion

To estimate the proportion of false positives in the main dataset, a sample of 30 systematically spaced 10 min periods was selected for analysis. For each of the 10 min sample periods, all hydrophones with detections were manually examined using a binary spectrogram display program (Moretti *et al.*, 2006). One of the sample periods was corrupted due to a system test being conducted by range technicians, and was removed from further analysis.

Experienced analysts classified the clicks and assigned click counts into one of the following seven categories: Md, delphinid, sperm whale, boat, pinger, other (which could include other beaked whale species), or "mixed" (i.e., when, in a given minute, Md clicks and other sound sources were not unambiguously distinguishable). Hence the false positive proportion might be between two extreme values, depending on whether all or none of the mixed detections is considered Md clicks. For the density estimates presented here both are used, leading to upper and lower bounds for density, with the best estimate of density likely lying somewhere in between these. The false positive proportion was estimated as the weighted mean of the proportion of false positives by sample period, with total number of clicks detected by sample period as weights. The variance of this weighted mean was estimated using the same procedure as described for the mean click rate variance.
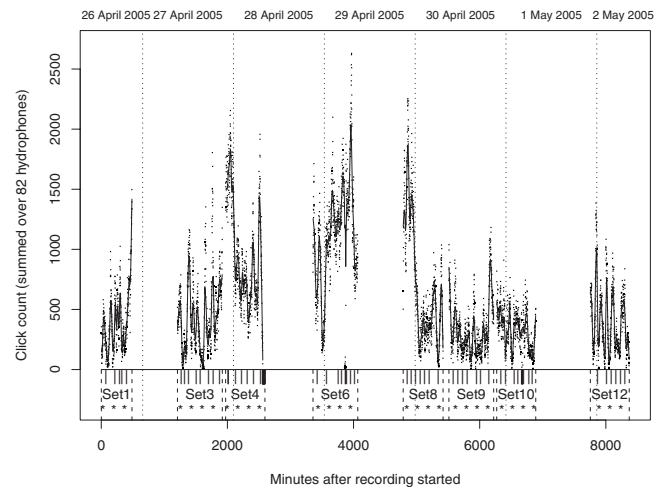

FIG. 2. Click counts per minute, summed over the 82 recording hydrophones, for the 6 day period of the primary dataset. Time is indexed as minutes since recording started. For operational reasons the data were divided into sets, and some sets (2, 5, 7 and 11) were not used. A standard lowess smooth of click counts over time is shown for the sets used. The small black vertical dashes ("|") are scattered minutes within the 8 sets used which were faulty and hence removed from the data. The sample periods used for the estimation of the false positive proportion are represented by "*". The limits of each day and set are represented by dotted and dashed lines, respectively.

C. Case study results

After removing minutes with problems, click counts were available for 4961 min in the primary dataset (Fig. 2).

Pooled over the 82 recording hydrophones, a total of 2 940 521 sounds classified as being beaked whale clicks were detected. Considering hydrophones as independent spatial replicates, the standard deviation of the total number of sounds was 161 373, resulting in a *CV* of 5.5%.

The estimated click rate was 0.407 (clicks/second), with a standard error of 0.040, resulting in a *CV* of 9.8%.

Overall, of the 160 302 sounds detected and considered to be beaked whale clicks during the sample periods used for the false positive proportion estimation, only 78 450 (corresponding to slightly under 50%) were unambiguously identified as Md clicks. The complement of the proportion of false positives, $1 - \hat{c}$, was estimated to be 0.549 ($se = 0.011$, $CV = 1.99\%$) or 0.489 ($se = 0.011$, $CV = 2.29\%$), depending on whether the mixed clicks are considered to be all or none from Md.

The estimated detection function is shown in Fig. 3. The maximum distance at which a click was detected was 6504 m. The estimated mean detection probability for clicks produced within 8 km of the hydrophones (clicks outside this buffer are assumed to have 0 detection probability) was 0.032, with an associated *CV* of 15.9%.

Using Eq. (3) the estimated Md density at AUTEC over the recording period was 25.3 or 22.5 animals per 1000 km^2 (with respectively 19.5% and 19.6% *CV*'s), depending on the $\hat{c}$ used. Assuming a log-normal distribution for the density estimate, the respective 95% confidence intervals for density are 17.3–36.9 and 15.4–32.9. Note that the contribution of n, $1 - \hat{c}$, $\hat{P}$, and $\hat{r}$ to the overall variance estimate on density is

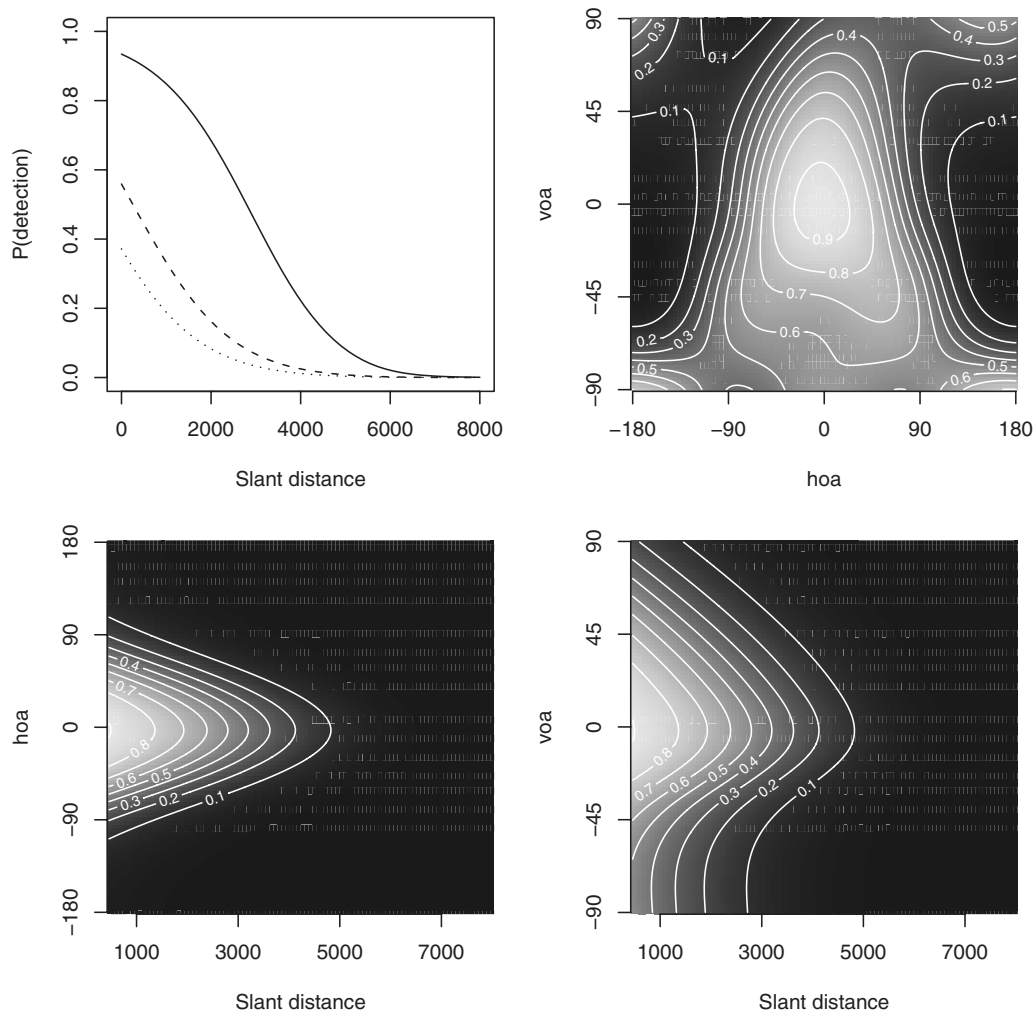Marques *et al.*: Cetacean density estimation from passive acoustics

FIG. 3. The estimated detection function. Partial plots (on the response scale) of the fitted smooths for a binomial GAM model with slant distance and a 2D smooth of *hoa* and *voa*. For the top left plot, the off-axis angles are fixed at 0, 45, and 90° (respectively the solid, dashed, and dotted lines). Remaining plots are two-dimensional representations of the smooths, where black and white represent respectively an estimated probability of detection of 0 and 1. Distance (top right panel) and angle not shown (bottom panels) are fixed respectively at 0 m and 0 °.

respectively, 8, 1, 66, and 25%.

IV. DISCUSSION

A. Assessment of case study

1. Comparison with previous work on beaked whales

The case study provided estimates of Md density at AU-TEC over a short time period of 6 days. Although true density is not known with certainty, the estimates obtained are consistent with other sources. The vast majority of confirmed beaked whale sightings in the Northern Bahamas are for Md (Claridge, 2006), but the presence of Cuvier's beaked whale (*Ziphius cavirostris*) and Gervais's Beaked Whale (*Mesoplodon europaeus*) have recently been noted at the range.[1] Clicks potentially from other species were treated as false positives, and hence the density estimate presented is valid for Md only.

Using a method based on isolating and counting groups of diving whales over a fixed time period, Moretti *et al.* (2006) obtained estimates of beaked whales at AUTEC using the same primary dataset as analyzed in this paper. Despite their method having some *ad hoc* components, requiring

strong assumptions about animal movement and an independent estimate of mean group size, the reported density estimates of 22.0 or 30.0 (depending on the estimate for mean group size used) Md per 1000 km² are consistent with the estimates obtained here. (Following a more recent, unpublished, revision of the their estimation process, the area covered by the survey was revised, leading to point estimates of 25.4 or 34.7 Md per 1000 km².)

The estimated detection function model (Fig. 3) was consistent with what was expected *a priori*: a decreasing function of distance and a decreasing function of the off-axis angles. This reflects the intuitive notions that the further away and off-axis a Md click is, the harder it is to detect. This also represents further evidence for the high directionality of beaked whale clicks, as has been demonstrated for Cuvier's beaked whale (Zimmer *et al.*, 2005).

The detection of a click seems to be slightly more likely if "below" the whale (negative *voa*) than "above" the whale (positive *voa*): an observation that deserves further investigation in characterizing the Md beam pattern.

At first sight, our results could seem to contradict those from Zimmer *et al.* (2008), which used the passive sonar

equation to model the maximum detection range of Cuvier's beaked whale clicks using a shallow hydrophone in the Mediterranean sea, and predicted a maximum detection range of 4 km. The pertinent parameters utilized were source level of 200 dB re 1 μPa at 1 m, 40 kHz center frequency and processing bandwidth, depth of 0 m, spectral noise level of 30 dB re 1 μPa^2 Hz^{-1}, absorption coefficient of 9.5 dB km^{-1}, and receiver threshold of 14 dB. The deeper hydrophones at AUTEC reduce transmission loss and masking noise, and the processing bandwidth at AUTEC was smaller. In addition, there is justification for assuming a higher maximum source level, since the measurements on which the 200 dB level is based is thought to be an underestimation (Zimmer *et al.*, 2005), and since source levels are known to vary considerably within and between individuals (Zimmer *et al.*, 2008). Johnson *et al.* (2006) reported center and −10 dB frequencies for Md measured at close range using acoustic tags of 38.3 and 26.6 kHz, and hence source levels of 210 and 200 dB re 1 μPa at 1 m at these two frequencies are plausible. Other pertinent parameters are 24 kHz processing bandwidth, depth of 1300 m (approximately mid-way between whale and hydrophone), spectral noise level of 26 dB re 1 μPa^2 Hz^{-1}, absorption coefficients of 8.9 and 5.1 dB km^{-1} at the two frequencies, and receiver threshold of 14 dB. These parameters provide maximum detection ranges of 5.7 km at 38.3 kHz and 7.6 km at 26.6 kHz. The largest observed detection range of a tagged animal in the case study was 6504 m; the above simplistic modeling exercise (e.g., assuming only spherical spreading) suggests that while the high frequency components of the click will be lost at such ranges, sufficient energy remains at lower frequencies to trigger the detector used. Supporting this, the mean received center frequency for clicks detected at 6.5 km range or greater was 29.7 kHz.

2. Assessment of detection probability estimation

Average detection probability was estimated in the case study using a sample of tagged whales for which it could be determined whether cues produced by the whale were detected on surrounding AUTEC hydrophones. However, the simple proportion of these cues detected could not be used directly to estimate average detection probability during the primary survey period for two reasons. First, the sample size was small (four whales over 13 dives), so it is unlikely that the distribution of potentially important covariates in the sample, such as distance from hydrophone, was representative of that over the 6 days of the primary survey. Second, the auxiliary (tag) data were collected at a different time period, when detection conditions were likely different.

The general solution to this problem is to construct a model that includes all the variables that affect detectability and for which the distribution of that variable differs between the auxiliary and the primary datasets. Model selection techniques could be used to select among a suite of candidate variables. Average detectability is then estimated for the correct values of these variables at the time of the primary survey. This approach requires that data are available on the variables affecting detectability both at the time of the auxiliary and primary surveys.

In the case study, because of the small sample size of tags, it was felt that variables indexing the whale's orientation and position relative to the hydrophones may differ between auxiliary and primary data. Both orientation and position were thought *a priori* to have an important influence on detectability. Since orientation and position were not known for cues received in the primary survey, a simulation approach was used to estimate the distribution of these variables during the primary survey. The simulation assumed that a whale's initial position was random with respect to hydrophone horizontal distance (because in the simulation whales were initially located at random within the AUTEC range), and that the dive behavior of the 21 dives from five whales for which data were available in the auxiliary dataset reflected the behavior of whales during the primary survey. The assumption that whale location is random relative to hydrophone seems largely reasonable, given that there were a large number of hydrophones throughout the study area. However, one may expect slightly fewer whales than predicted at larger horizontal distances, because some of the edge hydrophones cover areas that are too shallow for Md to dive in. Nevertheless, the overall influence of these "edge effects" is likely small.

There is no evidence against the assumption that the diving behavior recorded on the 21 dives is representative of that for whales in the primary survey, although this is currently untestable. Baird *et al.* (2008) reported no evidence of a tag-on effect, very similar dive times, and absence of diurnal patterns in dive cycles for Md (with tag durations of up to 72 h).

There are three possibly important variables affecting the detection function that were not considered in the case study. The first is potential differences between hydrophones, since it is possible (but not likely) that the distribution of sensitivities of hydrophones used in the auxiliary dataset is not representative of the 82 recording during the primary survey. The second relates to vertical directivity in the hydrophones, as raised by a reviewer. However, a preliminary analysis of the data revealed that including the vertical angle in the model produced inconsistent results. This is likely because most of the clicks within detection range are made within a narrow band of possible vertical angles with respect to the hydrophones, so any effect of vertical directivity was too negligible to estimate. Even if a strong effect existed, the results of the model reported here would have integrated it implicitly, provided the distribution of vertical angles in the DTag data was representative of the distribution in the 6 day dataset. The third, and potentially more important, is differences in ambient noise conditions, since tags can only be applied during calm weather. For example, mean wind speed, recorded at a location close to the AUTEC range, was 6.1 kn (range 1.1–13.8) during the 13 deep dives used for estimating the detection function, but 12.4 kn (range 2.5–20.8) during the primary survey period. Higher wind speed could reduce detectability, and hence the current estimate of average detection probability could be an overestimate. However the extent of the effect given the depth of the hydrophones remains to be determined. One potential approach would be to include variables such as wind speed, or ambient noise mea-

sured at the hydrophone, in the empirical detection function modeling. However, predicting detection probability for the primary survey period would then mean extrapolating significantly outside the range of the data since maximum wind speed was significantly higher in the primary survey period. An alternative would be to analytically model the effect of increased noise on the detection algorithm, but this is unlikely to be as robust as an empirical approach. A more attractive alternative is to take sound samples from the hydrophones during the primary survey period when no clicks were present and under a range of known wind speeds, use these to contaminate the hydrophone sound files used in the auxiliary study, and thereby empirically study the effect of increased noise on detectability. These approaches will be the subject of a future study.

The detection function was estimated considering only direct path detections. If multi-path detections were included in the click counts in the primary data set, then detection probability could have been underestimated and density correspondingly overestimated. While this is unlikely to have occurred, it might be worth to be investigated further.

It would have been better to have the primary dataset and the auxiliary dataset from the same time period. However, this was not possible because there were not readily available sound recordings for the entire range for the days in which there were DTags. Additionally, placement of DTags corresponded to events considerably spread in time and therefore the mean density estimated over these time periods would be less meaningful.

The method used to estimate probability of detection in the auxiliary dataset assumes that there are no false associations between clicks produced by the whale carrying the tag and clicks detected at the hydrophones. Conceivably, especially if there were other animals around the tagged whale, it is possible that some clicks produced by other animals were wrongly associated with detected clicks on the hydrophone. If this happened, one would overestimate the probability of detecting a click, leading to an underestimation of density. However, this false association seems unlikely: of all clicks detected at more than 4 km, 94.8% were at both *hoa* and *voa* of less than 30°; if these were false associations, one would expect that these correspond to a random sample of off-axis angles, rather than to mostly on-axis clicks. Another check would be to compare the received sound intensity at different frequencies within the 24–48 kHz range for the long-range detections to those predicted from a model of sound production and propagation.

3. Assessment of cue production rate estimation

The auxiliary dataset was also used in estimating click production rate, and hence the assumption is required that average click production rate during the time of the auxiliary survey is equal to that in the primary survey. Information reported in Baird *et al.* (2008) and Tyack *et al.* (2006) about the constancy of deep-dive cycles was fundamental to be confident that a single dive rate was adequate for both day and night periods.

Whenever the cue-count approach proposed here is used, there will be the requirement for an unbiased estimate of average cue production rate to convert cue density into animal density. The application of these methods therefore requires further fundamental research in this area, such as evaluation of cue rates and changes in cue rate as a function of time of day (e.g., Baumgartner and Fratantoni, 2008), year (e.g., Whitehead and Weilgart, 1990), or any other covariates (e.g., group size, season, bottom depth, geographic location, etc). Douglas *et al.* (2005) report that differences in worldwide vocal production (for sperm whales) means that estimates of abundance that are based on cue rates need to be based on data from the population of interest, rather than from another population or some global average.

Note that it is not sufficient to have estimates of the click rate (or inter-click interval) of diving whales from only during vocal parts of their dives—a quantity that is relatively easily obtained from hydrophone data. This will grossly overestimate the required cue production rate, since it ignores the silent time spent between dives. For animals that can spend long periods of time silent, long-term tags or (if feasible) focal follows from the surface are the only options, unless individuals can be tracked through individually recognizable vocalizations.

If cue rate is density dependent, these methods become harder to implement, and it becomes more important to estimate cue rate during the primary survey. If cue rate depends on local group size, then it is important to also estimate cue rate during the primary survey. Failing that, it is still possible to estimate mean cue rate provided one has estimates of the relationship between cue rate and group size, and also the distribution of group sizes during the primary survey.

4. Assessment of false positive proportion estimation

In studies where a manual analysis of the data provides a reliable "gold standard," false positive proportion is the easiest component to estimate. The variance of this component is also easily controllable by taking additional random samples for manual analysis until it is down to an acceptable level. In the case study, the contribution to the overall variance (around 1%) from this component was very small. This does not account for variance due to the specific human operator, and if this is believed to be a considerable source of variation one might consider having different operators repeating the task to quantify it. Here such a procedure was not deemed necessary.

Despite the method being general, it must be noted that values estimated under a given scenario are valid only under that scenario: the number of false positives is dependent on the number of sound sources producing false positives. As an example, everything else being the same, in an area where dolphins (a major source of false positives for the FFT detector used) were much more abundant than at AUTEC, one would expect a much larger proportion of false positives than what was found here.

About half of the sounds recorded as beaked whale clicks were estimated to be false positives in the case study. Such a high proportion was unexpected, although a fair proportion of false positives was expected *a priori*. This emphasizes the importance of accounting for false positives as part of density estimation. A second important lesson is that a

high false positive proportion is not necessarily a barrier to reliable density estimation: it is less important what the characteristics of the detector/classifier are than being able to accurately determine them. This is in contrast to other applications of cetacean detectors—for example, as part of a harm mitigation system—where a large number of false positives would lead to mitigation measures being triggered far more often than required.

The inability to unambiguously classify a proportion of the detected clicks led to increased uncertainty in the reported density estimates, with estimates changing by about 10% depending on how mixed clicks are accounted for.

B. Alternative approaches

1. Alternative methods for estimating P in a cue-counting framework

In the case study, the detection function was estimated based on the auxiliary tagging data. However, sound propagation modeling approaches like those used by Zimmer *et al.* (2008) or Stafford *et al.* (2007) might be alternatively used to obtain the detection function. Alternatively, one could implement broadcasting trials, by mimicking the animal's sound, to empirically estimate the detection function. This would naturally require a comprehensive understanding of the sound production and sound characteristics for the species under study. Ideally, multiple approaches would be used as a way to test the robustness of the results. In all of these cases, the detection function will have to be assumed to be representative of the detection function during the sampling period.

The advantages of using conventional distance sampling methods as described in Buckland *et al.* (2001) are as follows: (i) that the detection function is based on the distances to the detected cues, hence there is no need for additional data from tagging studies, and (ii) the detection function is necessarily valid for the conditions under which the data were collected.

For beaked whales at AUTEC such an approach would be hard to implement, because a given click is often not detected on more than one or two hydrophones, but in a setting where acoustic localization of cues is more feasible such an approach is recommended. Alternative methods of obtaining distances, such as using received sound levels or multi-path detections, may prove useful. Nonetheless, under such a setting there are two fundamental assumptions to estimate the detection function which were not required for the case study: (i) the probability of detecting a cue at distance zero must be 1 (or at least known) and (ii) the distribution of distances of all cues, detected or not, pooled across hydrophones, must be triangular (see Buckland *et al.*, 2001). There is an additional shortcoming for the use of conventional distance sampling methods: the distances required are those to the actual cues (i.e., excluding false positives); alternatively the detection function must be the same for true and false positives. The former seems hard to achieve, and the latter might be unreasonable under some scenarios.

As was described previously, had tags been applied to a large random sample of animals, then P could simply be estimated from the proportion of detected clicks, therefore avoiding the complex modeling exercise presented here. For the data at hand, such an estimate would be 0.036 (*CV* of 0.155), which is similar to the results from the GAM approach. Nevertheless, the simpler approach cannot be recommended in general for such a small number of dives because the assumptions underlying it are much more restrictive. Rather than having to be representative of the spread of the covariates included in a detection function model, the proportion of clicks detected by dive would have to be representative of the average P.

2. Alternative frameworks for estimating density from fixed passive acoustics

The framework presented here focused on estimating density from an analysis of cues, but other approaches can be envisaged based on detecting animals or groups of animals, and the potential for further research in this area is enormous. If animals can be located based on their vocalizations then the use of an approach similar to the "snapshot" point transect method (Buckland, 2006) could be envisaged. The method could be applied directly if animals could be located using single hydrophone data (e.g., Tiemann *et al.*, 2004); if multiple hydrophones were required for localization then additional development would be needed to account for the dependencies between hydrophones necessary to generate a distance. A major advantage of such an approach would be that it does not require an estimate of cue rate, and so is potentially feasible without any auxiliary data. Similar approaches could be used if groups could be isolated and localized, but additional information on mean group size would be required.

A disadvantage of snapshot type methods is that they require the counts to be discretized into intervals separated in time, so as to introduce statistical independence between samples. Alternative methods that explicitly account for the continuous nature of data acquisition but require information about animal movement exist in the historical literature (e.g., Skellam, 1958); however these need further development to account for issues such as imperfect detection.

In situations, such as for Md at AUTEC, where groups of diving animals can be isolated acoustically, even if not accurately localized, then there is the potential to use a cue counting approach based on counting group dive starts rather than individual clicks. At AUTEC, the probability of detecting a group dive start is 1, so there is no need to estimate the component P in Eq. (3). Similarly, false positives are relatively simple to identify. The cue rate required is the rate of dive initiations, and this could be obtained from time-depth-recording tags, or by tracking focal groups. Hence, such an approach seems potentially feasible and is the subject of ongoing research.

If detections of cues, animals, or groups on multiple hydrophones can be associated, then alternative density estimation methods may be possible that are based on mark-recapture, rather than distance sampling methods. Especially noteworthy are spatially explicit mark-recapture methods (Efford *et al.*, 2008). Mark-recapture methods have the advantage, over conventional distance sampling methods, of

not having to assume that the true distribution of animals with respect to samplers is known (Sec. II B), which is unrealistic if a small number of hydrophones is used. However, mark-recapture methods have their own assumptions, and it is not easy *a priori* to decide which will be more adequate in general.

V. CONCLUSIONS

The framework that has been laid out here is applicable to any study where the key components can be estimated: average detection probability, cue rate, and false positive proportion. Although the case study involved an array of 82 hydrophones over a relatively restricted area and time period, the same methods can be used on a far smaller number of hydrophones (with some loss of precision, obviously), or hydrophones distributed over a much larger area and sampling a longer time period.

Passive acoustic methods are potentially most useful for species that are poorly covered by visual methods but that produce loud, distinctive, predictable vocalizations. Towed acoustic (and visual) methods tend to be better where the goal is wide spatial coverage but narrow temporal coverage; fixed acoustic methods tend to be better where density estimates over long time periods are required. However, the optimal modality depends on many factors, including the study goals, target species, and current technology (the last of which is evolving rapidly). Optimal survey design is an important future research field, in terms of (i) optimal modality for a given situation, (ii) the trade-off between efforts on primary and auxiliary surveys (if required), and (iii) the number and placement of sensors.

Although the focus of this paper has been on estimating cetacean abundance using underwater acoustic data, there is no conceptual extra step required for implementing it with sound recording devices in terrestrial environments. Potential examples include forest surveys of songbirds, territorial monkey species such as gibbons, and elephants (e.g., Payne *et al.*, 2003). Fixed passive acoustics methods are potentially invaluable for monitoring density over time at multiple sites without requiring a large amount of human resources, and for this reason will probably become common practice in the future.

Regarding monitoring of cetaceans, to date most studies have used acoustic data as an index of abundance over time or space (e.g., Stafford *et al.*, 2007). This requires strong (and largely untestable) assumptions that the factors relating number of calls detected to density remain constant (see Anderson, 2001, 2003, for a discussion of these issues). Here, such approaches are improved upon by allowing estimation of these factors, leading to inferences about density.

Mellinger *et al.* (2007b) urged the need for the development of statistical methods for estimating cetacean populations acoustically. This paper shows that the foundations for such methods are laid, and hopes to contribute to their widespread use.

[1]This information was provided by Diane E. Claridge, from the Bahamas Marine Mammal Research Organisation.

Anderson, D. R. (**2001**). "The need to get the basics right in wildlife field studies," Wild. Soc. Bull. **29**, 1294–1297.

Anderson, D. R. (**2003**). "Response to Engeman: Index values rarely constitute reliable information," Wild. Soc. Bull. **31**, 288–291.

Au, W. W. L., Mobley, J., Burgess, W. C., Lammers, M. O., and Nachtigall, P. E. (**2000**). "Seasonal and diurnal trends of chorusing humpback whales wintering in waters off western Maui," Marine Mammal Sci. **16**, 530–544.

Aubauer, R., Lammers, M. O., and Au, W. W. L. (**2000**). "One-hydrophone method of estimating distance and depth of phonating dolphins in shallow water," J. Acoust. Soc. Am. **107**, 2744–2749.

Baird, R. W., Webster, D. L., Schorr, G. S., Mcsweeney, D. J., and Barlow, J. (**2008**). "Diel variation in beaked whale diving behavior," Marine Mammal Sci. **24**, 630–642.

Barlow, J., and Gisiner, R. (**2006**). "Mitigating, monitoring and assessing the effects of anthropogenic sound on beaked whales," J. Cetacean Res. Manage. **7**, 239–249.

Barlow, J., and Taylor, B. (**2005**). "Estimates of sperm whale abundance in the northeastern temperate Pacific from a combined acoustic and visual survey," Marine Mammal Sci. **21**, 429–445.

Barlow, J., Ferguson, M. C., Perrin, W. F., Ballance, L., Gerrodette, T., Joyce, G., Macleod, C. D., Mullin, K., Palka, D. L., and Waring, G. (**2006**). "Abundance and densities of beaked and bottlenose whales (family Ziphiidae)," J. Cetacean Res. Manage. **7**, 263–270.

Baumgartner, M. F., and Fratantoni, D. M. (**2008**). "Diel periodicity in both sei whale vocalization rates and the vertical migration of their copepod prey observed from ocean gliders," Limnol. Oceanogr. **53**, 2197–2209.

Buckland, S. T. (**2006**). "Point transect surveys for songbirds: Robust methodologies," Auk **123**, 345–345.

Buckland, S. T., Anderson, D. R., Burnham, K. P., Laake, J. L., Borchers, D. L., and Thomas, L. (**2001**). *Introduction to Distance Sampling: Estimating Abundance of Biological Populations* (Oxford University Press, Oxford).

Buckland, S. T., Anderson, D. R., Burnham, K. P., Laake, J. L., Borchers, D., and Thomas, L. (**2004**). *Advanced Distance Sampling* (Oxford University Press, Oxford).

Charif, R. A., Clapham, P. J., and Clark, C. W. (**2001**). "Acoustic detections of singing humpback whales in deep waters off the British Isles," Marine Mammal Sci. **17**, 751–768.

Claridge, D. E. (**2006**). "Fine-scale distribution and habitat selection of beaked whales," MS thesis, University of Aberdeen, Scotland.

DiMarzio, N., Moretti, D., Ward, J., Morrissey, R., Jarvis, S., Izzi, A. M., Johnson, M., Tyack, P., and Hansen, A. (**2008**). "Passive acoustic measurement of dive vocal behavior and group size of Blainville's beaked

whale (*Mesoplodon densirostris*) in the Tongue of the Ocean (TOTO)," Can. Acoust. **36**, 166–173.

Douglas, L. A., Dawson, S. M., and Jaquet, N. (**2005**). "Click rates and silences of sperm whales at Kaikoura, New Zealand," J. Acoust. Soc. Am. **118**, 523–529.

Efford, M. G., Borchers, D. L., and Byrom, A. E. (**2008**). "Density estimation by spatially explicit capture-recapture: Likelihood-based methods," in *Modeling Demographic Processes in Marked Populations*, edited by D. L. Thomson, E. G. Cooch, and M. J. Conroy (Springer, New York), pp. 255–269.

Efron, B., and Tibshirani, R. (**1993**). *An Introduction to the Bootstrap* (Chapman and Hall, London/CRC, New York).

Freitag, L. E., and Tyack, P. L. (**1993**). "Passive acoustic localization of the Atlantic bottlenose dolphin using whistles and echolocation clicks," J. Acoust. Soc. Am. **93**, 2197–2205.

Gatz, D. F., and Smith, L. (**1995**). "The standard error of a weighted mean concentration—I. Bootstrapping vs other methods," Atmos. Environ. **29**, 1185–1193.

Hastie, G. D., Swift, R. J., Gordon, J. C., Slesser, G., and Turrell, W. R. (**2003**). "Sperm whale distribution and seasonal density in the Faroe Shetland Channel," J. Cetacean Res. Manage. **5**, 247–252.

Johnson, M. P., and Tyack, P. L. (**2003**). "A digital acoustic recording tag for measuring the response of wild marine mammals to sound," IEEE J. Ocean. Eng. **28**, 3–12.

Johnson, M., Madsen, P. T., Zimmer, W. M. X., de Soto, N. A., and Tyack, P. L. (**2004**). "Beaked whales echolocate on prey," Proc. R. Soc. London, Ser. B **271**, S383–S386.

Johnson, M., Madsen, P. T., Zimmer, W. M. X., de Soto, N. A., and Tyack, P. L. (**2006**). "Foraging Blainville's beaked whales (*Mesoplodon densirostris*) produce distinct click types matched to different phases of echolocation," J. Exp. Biol. **209**, 5038–5050.

Lewis, T., Gillespie, D., Lacey, C., Matthews, J., Danbolt, M., Leaper, R., McLanaghan, R., and Moscrop, A. (**2007**). "Sperm whale abundance estimates from acoustic surveys of the Ionian Sea and Straits of Sicily in 2003," J. Mar. Biol. Assoc. U.K. **87**, 353–357.

McDonald, M. A., and Fox, C. G. (**1999**). "Passive acoustic methods applied to fin whale population density estimation," J. Acoust. Soc. Am. **105**, 2643–2651.

Mellinger, D. K., Nieukirk, S. L., Matsumoto, H., Heimlich, S. L., Dziak, R. P., Haxel, J., Fowler, M., Meinig, C., and Miller, H. V. (**2007a**). "Seasonal occurrence of North Atlantic right whale (*Eubalaena glacialis*) vocalizations at two sites on the Scotian Shelf," Marine Mammal Sci. **23**, 856–867.

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

Moore, S. E., Stafford, K. M., Mellinger, D. K., and Hildebrand, J. A. (**2006**). "Listening for large whales in the offshore waters of Alaska," BioScience **56**, 49–55.

Moretti, D., DiMarzio, N., Morrissey, R., Ward, J., and Jarvis, S. (**2006**). "Estimating the density of Blainville's beaked whale (*Mesoplodon densirostris*) in the Tongue of the Ocean (TOTO) using passive acoustics," in Proceedings of the Oceans'06 MTS/IEEE-Boston, Boston, MA.

Morrissey, R. P., Ward, J., DiMarzio, N., Jarvis, S., and Moretti, D. J. (**2006**). "Passive acoustic detection and localization of sperm whales (*Physeter macrocephalus*) in the tongue of the ocean," Appl. Acoust. **67**, 1091–1105.

Payne, K. B., Thompson, M., and Kramer, L. (**2003**). "Elephant calling patterns as indicators of group size and composition: The basis for an acoustic monitoring system," Afr. J. Ecol. **41**, 99–107.

Powell, L. A. (**2007**). "Approximating variance of demographic parameters using the delta method: A reference for avian biologists," Condor **109**, 949–954.

R Development Core Team (**2008**). "R: A language and environment for statistical computing," URL: http://www.R-project.org (Last viewed March 2009).

Seber, G. A. F. (**1982**). *The Estimation of Animal Abundance*, 2nd ed. (Griffin, London).

Skellam, J. G. (**1958**). "The mathematical foundations underlying the use of line transects in animal ecology," Biometrics **14**, 385–400.

Stafford, K., Mellinger, D., Moore, S., and Fox, C. (**2007**). "Seasonal variability and detection range modeling of baleen whale calls in the Gulf of Alaska, 1999–2002," J. Acoust. Soc. Am. **12**, 3378–3390.

Thomas, L., Laake, J. L., Strindberg, S., Marques, F. F. C., Buckland, S. T., Borchers, D. L., Anderson, D. R., Burnham, K. P., Hedley, S. L., Pollard, J. H., Bishop, J. R. B., and Marques, T. A. (**2006**). "Distance 5.0," Research Unit for Wildlife Population Assessment, University of St. Andrews, UK. http://www.ruwpa.st-and.ac.uk/distance/(Last viewed March, 2009).

Tiemann, C. O., Porter, M. B., and Frazer, L. N. (**2004**). "Localization of marine mammals near Hawaii using an acoustic propagation model," J. Acoust. Soc. Am. **115**, 2834–2843.

Tyack, P. L., Johnson, M., Soto, N. A., Sturlese, A., and Madsen, P. T. (**2006**). "Extreme diving behaviour of beaked whales," J. Exp. Biol. **209**, 4238–4253.

Ward, J., Morrissey, R., Moretti, D., DiMarzio, N., Jarvis, S., Johnson, M., Tyack, P., and White, C. (**2008**). "Passive acoustic detection and localization of *Mesoplodon densirostris* (Blainville's beaked whale) vocalizations using distributed bottom-mounted hydrophones in conjunction with a digital tag (DTag) recording," Can. Acoust. **36**, 60–66.

Whitehead, H., and Weilgart, L. (**1990**). "Click rates from sperm whales," J. Acoust. Soc. Am. **87**, 1798–1806.

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

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

Zimmer, W. M. X., Johnson, M. P., Madsen, P. T., and Tyack, P. L. (**2005**). "Echolocation clicks of free-ranging Cuvier's beaked whales (*Ziphius cavirostris*)," J. Acoust. Soc. Am. **117**, 3919–3927.

Zimmer, W. M., Harwood, J., Tyack, P. L., Johnson, M. P., and Madsen, P. T. (**2008**). "Passive acoustic detection of deep diving beaked whales," J. Acoust. Soc. Am. **124**, 2823–2832.