

Passive acoustic detection of deep-diving beaked whales

Walter M. X. Zimmer^{a)}

NATO Undersea Research Centre, Viale San Bartolomeo 400, 19138 La Spezia, Italy

John Harwood

Centre for Research Into Ecological and Environmental Modelling, University of St Andrews,
Fife KY16 9LZ, United Kingdom

Peter L. Tyack and Mark P. Johnson

Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02453

Peter T. Madsen

Zoophysiology, Department of Biological Sciences, University of Aarhus, 6000 Aarhus, Denmark

(Received 23 April 2008; revised 18 August 2008; accepted 21 August 2008)

Beaked whales can remain submerged for an hour or more and are difficult to sight when they come to the surface to breathe. Passive acoustic detection (PAD) not only complements traditional visual-based methods for detecting these species but also can be more effective because beaked whales produce clicks regularly to echolocate on prey during deep foraging dives. The effectiveness of PAD for beaked whales depends not only on the acoustic behavior and output of the animals but also on environmental conditions and the quality of the passive sonar implemented. A primary constraint on the range at which beaked whale clicks can be detected involves their high frequencies, which attenuate rapidly, resulting in limited ranges of detection, especially in adverse environmental conditions. Given current knowledge of source parameters and in good conditions, for example, with a wind speed of 2 m/s, a receiver close to the surface should be able to detect acoustically Cuvier's beaked whales with a high probability at distances up to 0.7 km, provided the listening duration exceeds the deep dive interval, about 2.5 h on average. Detection ranges beyond 4 km are unlikely and would require low ambient noise or special sound propagation conditions.

© 2008 Acoustical Society of America. [DOI: 10.1121/1.2988277]

PACS number(s): 43.30.Sf [WWA]

Pages: 2823–2832

I. INTRODUCTION

Recent findings that strandings of beaked whales may be caused by human activities, such as sonar exercises (Cox *et al.*, 2006), have driven a need for improved methods to detect beaked whales. Such methods are important both for planning, so that areas with high concentrations of beaked whales can be avoided, and for mitigation to reduce the risk that sensitive animals are near ongoing exercises.

Cetaceans can be detected visually or acoustically with varying success depending on environmental conditions and their behavior (Buckland *et al.*, 2001; Barlow and Taylor, 2005). Visual detection of deep-diving cetacean species is limited in the best of sighting conditions by the short duration of their surfacing compared to their dive time. Although some species, such as sperm whales, spend considerable time visible at the surface between their deep dives and have visible blows, beaked whales only surface for a few seconds at a time and produce nearly invisible blows. The probability of visual detection for beaked whales is so low that it is of little practical use for monitoring purposes (Barlow, 1999; Barlow and Gisiner 2006). Barlow and Taylor (2005) and Barlow and Gisiner (2006) considered acoustic detection to be a more promising approach.

Whales can be detected acoustically either by using whale-finding sonar to listen for echoes reflected from the whale (active acoustic detection) or by listening for sounds emitted by the whales [passive acoustic detection (PAD)]. Whale-finding sonars do not require the whale to make a sound but to be effective, they require substantial sound energy to obtain detectable echoes. This is because the air in the lungs of a whale is compressed during a deep dive, thus reducing the whale's target strength significantly at depth. While some success in detecting baleen whales with active sonar has been reported (Lucifredi and Stein, 2007), the feasibility of active acoustic detection has not yet been demonstrated for deep-diving whales and the increased sound energy required to detect these species may generate additional risks.

PAD on the other hand is based on listening to the acoustic output from whales. Deep-diving toothed whales are known to employ echolocation to find and capture prey and are therefore acoustically very active while foraging. For example, sperm whales emit over 2500 clicks per foraging dive (Madsen *et al.*, 2002). As the spectra of sperm whale clicks reach into the audio band, they are easily detected and therefore a valuable resource for underwater bioacoustic researchers (e.g., Watkins, 1980; Gordon, 1987; Møhl *et al.*, 2000; Wahlberg, 2002; Zimmer *et al.*, 2003; Drouot *et al.*, 2004; Mellinger *et al.*, 2004; Thode, 2004; Teloni, 2005). Beaked whales also produce thousands of clicks per foraging dive

^{a)}Electronic mail: zimmer@nurc.nato.int

(Johnson *et al.*, 2004, 2006; Madsen *et al.*, 2005; Zimmer *et al.*, 2005a), although most of the energy of these clicks lies above the frequencies humans can hear.

The successful application of PAD requires both appropriate technology and appropriate operational concepts. While the impact of technology or system parameters (e.g., self-noise, array gain, processing bandwidth, and gain) is easy to assess, the impact of operational factors is more difficult to quantify because they depend on the behavior of the whales (where, when, and how they echolocate) as well as on environmental characteristics that vary in time and space, influencing the design and performance of a PAD device. For example, the hydrophone depth required to maximize range and detection probability will influence the kinds of platforms that can be used to deploy it. Other operational issues include the mobility of the platform (whether the hydrophones should be moored, allowed to float, or can be towed from a ship) and the number of sensors required for obtaining the required success rate and confidence. The design of the sonar system will also depend on the objective of PAD: a system for risk mitigation, in which failure to detect a whale that is present constitutes an error, will have very different requirements than one designed for abundance estimation.

This paper outlines the requirements, and the problems, of using PAD to detect deep-diving beaked whales. In particular, we show that estimating the detection function (i.e., the probability of detecting an acoustically active deep-diving whale as function of its distance from the receiver) is a highly informative way of characterizing a PAD system.

II. METHOD AND RESULTS

Beaked whales emit two types of echolocation clicks: so-called regular clicks, suited for searching for prey items, and weaker buzz clicks used in the final stages of prey capture (Madsen *et al.*, 2005; Johnson *et al.*, 2006). In the following analysis we only consider regular clicks because they are 15 dB or more stronger than buzz clicks (Madsen *et al.*, 2002, 2005), and hence much more likely to be detected.

A. Sonar equation

The performance of a PAD system can be characterized by the passive sonar equation, which is a standardized way to assess whether or not a sound will be received with sufficient signal-to-noise ratio (SNR) for detection and classification (Lurton, 2002). The passive sonar equation combines source and environmental parameters in decibels and indicates that a sound will only be detected if the received SNR, after any filtering within the passive sonar, exceeds a predetermined receiver threshold (RT), i.e., if

$$\text{SNR} > \text{RT}. \quad (1)$$

Detections are not always correct because they are the result of a decision process. Incorrect detections occur when the target signal is not present but a transient in the background noise results in the received level exceeding RT. These events are called false alarms in sonar terminology or type I errors in statistics. The percentage of correct and wrong decisions is influenced by the choice of RT. The

higher the value of RT, the fewer false alarms will occur, but at the same time an increasing number of weaker target signals may be missed (i.e., the proportion of type II statistical errors will increase) and the power of the detector or the probability of a correct decision will decrease.

There are a variety of approaches for selecting RT, depending on available *a priori* information and constraints. If we opt for the Neyman–Pearson criterion, RT depends on the acceptable false alarm rate P_{FA} of the detector (Lurton, 2002). That is, RT is chosen in such a way that the passive sonar system generates (false) detections in the absence of whale clicks with a predefined probability P_{FA} . If background noise is assumed to be Rayleigh distributed, RT is given by

$$\text{RT} = 10 \log(-2 \ln(P_{\text{FA}})). \quad (2)$$

In cases where the distribution of background noise is known to take a different form, RT can be derived empirically.

The detection probability P_{det} is obtained by statistical analysis of the predicted or achieved detections and depends, in general, on the relative geometry between sound source and the hydrophone of the receiver, here expressed as slant range R :

$$P_{\text{det}}(R) = \Pr\{\text{SNR}(R) > \text{RT}\} = \int_{\text{RT}}^{\infty} w(\text{SNR}) d\text{SNR}, \quad (3)$$

where $w(\text{SNR})$ is the probability density function (PDF) of the SNR at range R .

Equations (2) and (3) show that P_{det} and P_{FA} are related through RT. This relationship is often represented in the form of receiver operating characteristics (Lurton, 2002), which will also vary with the value of R . While the probability of false alarm is uniquely characterized by the noise statistics, the probability of detection depends on the properties of $w(\text{SNR})$, which reflects our uncertainty about the exact received level of the whale sound.

The SNR at the detector input depends on a variety of parameters, as shown in the following equation:

$$\text{SNR} = \text{SL} - \text{DL}(\vartheta) - \text{TL}(R) - \text{NL}_0 + 10 \log B + \text{AG} + \text{PG}, \quad (4)$$

where SL is the (on-axis) source level of the sound source, DL(ϑ) is the attenuation of the source level as a consequence of the off-axis angle ϑ (i.e., the angle between the acoustic axis of the source and the direction from the whale to the hydrophone),¹ TL(R) is the range-dependent transmission loss, NL₀ is the spectral noise level of any masking noise at the receiver position, B is the processing bandwidth, AG is the gain of the hydrophone array, and PG is the processing gain of the receiver. All of these parameters are measured in dB, except SL which is measured in dB_{rms} re 1 μPa re 1 m,² NL₀ which is measured in dB_{rms} re 1 $\mu\text{Pa}^2 \text{Hz}^{-1}$, and B which is measured in Hz.

TL is a key term in the sonar equation. For relatively small values of R and constant sound speed it can be estimated by

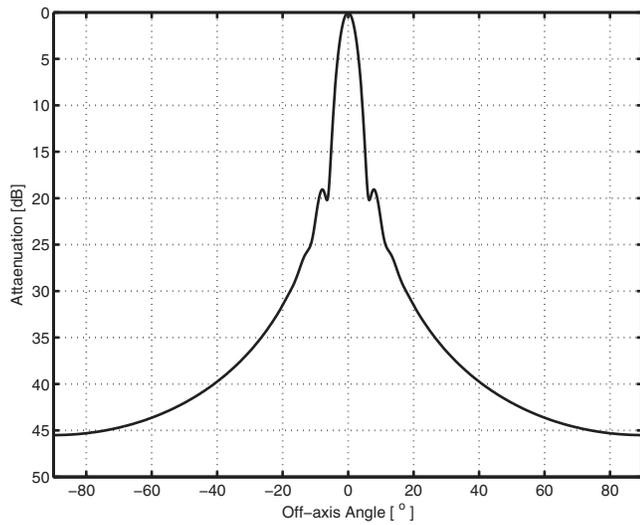


FIG. 1. Off-axis dependent attenuation of the echolocation signal of Cuvier's beaked whale adapted from Fig. 7 of Zimmer *et al.* (2005b).

$$TL(R) = 20 \log(R) + \alpha_f(R/1000), \quad (5)$$

where range R is expressed in m and α_f is the frequency dependent absorption coefficient in dB km^{-1} .

Clicks of Cuvier's and Blainville's beaked whales are centered around 40 kHz (Johnson *et al.* 2004, 2006; Zimmer *et al.*, 2005a), for which the absorption coefficient is about 9.5 dB km^{-1} (Lurton, 2002).

We assume that local ambient noise is mainly generated by surface waves and can therefore be described by the Knudsen curves [Lurton, 2002, Eqs. (4.12)]. This noise will decrease as it propagates from the surface to the bottom. Thus the SNR, and therefore P_{det} , will increase with the receiver depth [Lurton, 2002, Eq. (4.41)].

While typical values can be used for most of the quantities in the passive sonar equation, the off-axis dependent attenuation of the source level, or off-axis loss, $DL(\vartheta)$ of echolocating whales needs to be evaluated explicitly. Here, we consider that beaked whales have a broadband biosonar system (Zimmer *et al.*, 2005a) and that we have a rough knowledge of the transmission beam pattern, as presented in Fig. 1. The beam pattern behind the animal ($|\vartheta| > 90^\circ$) is assumed constant at the maximum attenuation level.

B. Basic detection function for a single click

A basic detection function for a single beaked whale click can be determined from Eq. (3), provided that suitable values for the constants of the sonar equation are available to estimate the SNR [Eq. (4)]. In order to illustrate this process we have used the values shown in Table I. In Sec. III we describe the sensitivity of our results to these values. The spectral noise level in Table I corresponds to a light breeze or a wind speed of 2.5 m s^{-1} for a near surface hydrophone, and the receiver threshold of 14 dB is for Rayleigh distributed ambient noise equivalent to $P_{\text{FA}} = 3.5 \times 10^{-6}$. This is equivalent to accepting a false detection about every minute for a processing window of 0.2 ms.

The relationship between SNR, R , and ϑ is shown in Fig. 2; the white threshold line divides the SNR map into

TABLE I. Constants of the sonar equation used to construct Figs. 2 and 3. Values were chosen to be characteristic of beaked whales and are consistent with Johnson *et al.* (2004) and Zimmer *et al.* (2005a). Effects of depth on the spectral noise level are not included.

Source level	200 dB_{rms} re 1 μPa at 1 m
Center frequency	40 kHz
Processing bandwidth	40 kHz
Spectral noise level	30 dB re 1 $\mu\text{Pa}^2/\text{Hz}$
Array gain	0 dB
Processing gain	0 dB
Sound attenuation	9.5 dB/km
Receiver threshold	14 dB

two parts: to the left of this line, the SNR exceeds RT and detection is possible, but to the right the SNR is too low for a successful detection.

Figure 2 was constructed on the assumption that nothing is known about the distribution of the aspect of an echolocating beaked whale with respect to the receiver. Nevertheless, one can deduce some important results from Fig. 2. In particular, for the values in Table I, a single beaked whale is very unlikely to be detected at ranges greater than 4 km because the SNR for all off-axis angles is below the threshold for the assumed propagation condition and source parameters. Using similar arguments, we can conclude that all clicks will be detected for ranges below about 0.7 km as the SNR for all off-axis angles is above the threshold. Between these limits, detectability depends on the probability that the whale is clicking in the direction of the receiver. Specific detection probabilities can be obtained if the orientation statistics for a particular location of animal and receiver are available and all other terms of SNR [Eq. (4)] are known because Eq. (1) is then equivalent to

$$P_{\text{det}}(R_H, h) = \int_{\text{SNR}(\vartheta) > \text{RT}} w(\vartheta) d\vartheta, \quad (6)$$

where R_H is the horizontal displacement of the whale from the hydrophone, h is the depth of the hydrophone, and $w(\vartheta)$

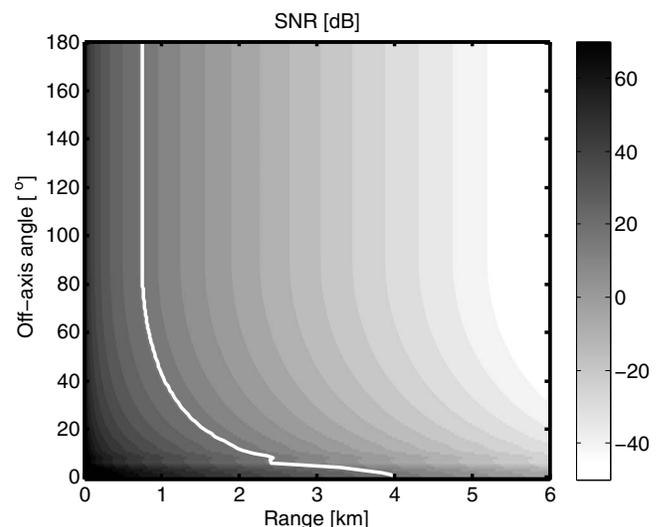


FIG. 2. SNR as a function of the off-axis angle of the source and its slant range. The white line indicates the threshold $\text{RT} = 14 \text{ dB}$.

is the PDF of the off-axis angle and the integral runs over all off-axis angles for which the SNR exceeds RT.

C. Off-axis angle statistics

The actual off-axis angle depends on the location of the whale relative to the hydrophone and its orientation while echolocating. If we assume that the whale orientation is not correlated with the location of the hydrophone, then we may choose any coordinate system, for example, [E, N, Z], where the x -axis corresponds to east (E), the y -axis to north (N), and the z -axis to the zenith (Z). The acoustic axis or direction vector of the whale when making a particular click is given by

$$\hat{c} = \begin{pmatrix} c_E \\ c_N \\ c_Z \end{pmatrix} = \begin{pmatrix} \sin \gamma \cos \beta \\ \cos \gamma \cos \beta \\ \sin \beta \end{pmatrix}, \quad (7)$$

where γ is the whale's heading (measured from north to east) and β represents its pitch (positive pitch is measured upward). The direction of the hydrophone, as seen by the whale, is

$$\hat{s} = \begin{pmatrix} s_E \\ s_N \\ s_Z \end{pmatrix} = \begin{pmatrix} \cos \eta \\ 0 \\ \sin \eta \end{pmatrix}, \quad (8)$$

where η is the elevation angle of the hydrophone from the whale, which is given by

$$\eta = \tan^{-1} \left(\frac{d-h}{R_H} \right), \quad (9)$$

where d is the depth of the whale and h is the depth of the hydrophone. In Eq. (8), the hydrophone is modeled as being due east of the whale, but there is no loss of generality as the whale can be in an arbitrary orientation via Eq. (7).

The off-axis angle ϑ (i.e., the angle between the whale's acoustic axis and the hydrophone direction) is then given by

$$\cos \vartheta = \hat{c}^T \hat{s}, \quad (10)$$

that is,

$$\vartheta(\gamma, \beta, \eta) = \cos^{-1}(\sin \gamma \cos \beta \cos \eta + \sin \beta \sin \eta), \quad (11)$$

and slant range R is simply determined from the relationship

$$R = \sqrt{R_H^2 + (d-h)^2}. \quad (12)$$

The way in which variations in off-axis angle affect the form of the detection function will depend on how a detection is defined. Here we consider a situation in which we want to determine the probability that at least one click will be detected during an entire dive sequence. In this case, all we need to know is the minimum off-axis angle adopted by a whale within a dive:

$$\vartheta_{\min}(\beta, \eta) = \min_{\text{dive}} \{ \cos^{-1}(\sin \gamma \cos \beta \cos \eta + \sin \beta \sin \eta) \}. \quad (13)$$

If the SNR associated with this orientation does not exceed RT, then the whale cannot be detected. If the SNR does

exceed RT, then other clicks within the dive may also be detected, but this will not affect the dive P_{det} . Ideally, statistical distributions for γ , β , and η should be obtained from direct observations of beaked whale dives. However, although some data for beaked whales are available (Tyack *et al.*, 2006; Johnson *et al.*, 2006), these are limited to a very small sample of animals. Instead, we assumed distributions for the parameters that would be expected for a beaked whale foraging on prey that are uniformly distributed in a horizontal layer around the predator. We recognize that beaked whales may feed on more than one layer, and in Sec. III we consider the sensitivity of our results to this simplification.

The distribution of whale pitch was modeled as a circular normal, or von Mises, distribution (Fisher, 1993):

$$w(\beta) = \frac{1}{2\pi I_0(\kappa)} \exp(\kappa \cos(\beta)), \quad (14)$$

where κ is a parameter related to the variance of the distribution by $\text{var}(\beta) = 1 - I_1(\kappa)^2 / I_0(\kappa)^2$ and $I_j(x)$ is the modified Bessel function of order j . For the present analysis a value of $\kappa=6$ was chosen, equivalent to a standard deviation of about 23° , which means that only a few clicks are oriented straight up or down, consistent with foraging in a horizontal layer. The distribution of heading within a dive was also modeled as a circular normal distribution relative to a mean whale heading γ_0 with the same standard deviation as for the pitch distribution, i.e., $\kappa=6$.

The distribution of the hydrophone elevation relative to the whale was modeled as a function of hydrophone depth and horizontal distance according to Eq. (9). We assumed that whale depth during a dive was normally distributed around a mean foraging depth d_0 .

$$w(d) = \frac{1}{\sqrt{2\pi}\sigma_d} \exp\left(-0.5\left(\frac{d-d_0}{\sigma_d}\right)^2\right). \quad (15)$$

For the present analysis d_0 was set to 720 m, with a standard deviation $\sigma_d=50$ m, and h to 100 m. Later, we examine the consequences of choosing different values for these parameters. We considered horizontal ranges from 0 to 10 km.

To obtain the PDF of the off-axis angle for a given scenario, or geometry, the assumed distributions $w(\gamma)$, $w(\beta)$, and $w(\eta)$ were transformed into $w(\vartheta)$. Although this could have been done analytically, it was straightforward to use a Monte Carlo simulation. In order to simulate a single dive, we drew 4000 random values from each of the distributions defined above. These were then transformed into a vector of random off-axis angles. This vector was used to simulate a whale's orientation each time it clicked during that dive. Although the pitch, heading, and elevation of a whale when a click is generated are likely to be highly correlated with its orientation on previous and subsequent clicks, our concern is only whether or not any of the clicks made during a dive can be detected. In this circumstance, the assumption of independence for the three PDFs is justified.

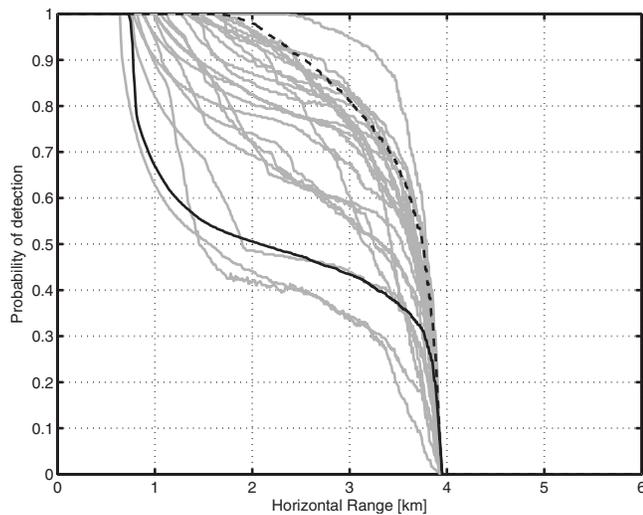


FIG. 3. Detection functions for a beaked whale echolocating at a depth of 720 m with a receiver depth of 100 m. The solid black curve corresponds to a simple simulation with normally distributed heading, pitch, and elevation. The gray lines show the probability of detection for clicks generated during 23 measured dives made by six Cuvier's beaked whale (*Zc*). The dashed black curve shows the effect of allowing a simulated whale to reverse the direction of its travel periodically.

D. Detection function for an individual dive

The process of dive simulation was repeated for all mean headings relative to the hydrophone (i.e., we rotated the mean whale heading γ_0 around the circle in steps of 0.1°) to obtain a range-dependent PDF for the off-axis angle. The detection function was then estimated by integrating this PDF according to Eq. (6). The result of this integration is shown by the solid black curve in Fig. 3.

This detection function has some properties in common with the one shown in Fig. 2. For example, all dives made within 0.7 km horizontal separation of the hydrophone from the whale are certain to be detected and no dives made beyond 4.0 km will be detected, for the source and environmental parameters used here, because absorption limits the detectability even of on-axis clicks. However, the distribution of detection probabilities within this range will vary, depending on the actual movements of the whale.

For comparison, we have also included 23 detection functions (shown in gray) for real dive profiles for six Cuvier's beaked whales (*Ziphius cavirostris*) in Fig. 3. For this data set the random pitch, heading, and depth distributions were replaced by actual measurements collected by digital acoustic recording tags (DTAG; Johnson and Tyack 2003) in the Ligurian Sea during 2003 and 2004 (Tyack *et al.*, 2006), that is, the detection functions were estimated in analogy to the simulated dives with the exception that whale pitch, heading, and depth were not modeled according to Eqs. (14) and (15) but obtained from the DTAG. The dashed black line in Fig. 3 was obtained by replacing the monomodal heading PDF [Eq. (14)], as used for the solid black line in Fig. 3, by a dual-modal PDF simulating a reversal of the swimming direction with a probability of 50%. This would represent a feeding behavior of directional swimming back and forth, for example, along a certain bathymetric contour. Both the solid and the dashed line in Fig. 3 characterize the boundaries of

the simulated detection functions. As seen from Fig. 3, most detection functions based on DTAG data fall within these limits and the few exceptions (e.g., to the left of the solid line) are due to residual mismatch in modeling parameters, for example, deeper foraging depth would result in shorter detection ranges.

III. DISCUSSION

In this paper, we have used diving and acoustic data of beaked whales along with theoretical considerations based on the sonar equation to explore the potential and limitations of PAD for deep-diving echolocating beaked whales. While the sonar equation is the tool of choice for assessing the performance of human sonar systems, it is a simplified mathematical description of a complicated problem. We simplified this system even further by analyzing only a single, but characteristic, parameter set shown in Table I. However, the sensitivity of our results to variation in these parameters can, in most cases, be evaluated in a straightforward way.

A. Importance of the components of the sonar equation

Many of the components of the sonar equation have a simple additive effect on the SNR associated with a whale click and therefore on the probability that it will be detected. Thus the effect of a 10 dB reduction in source level, a 10 dB increase in ambient noise at the depth of the PAD hydrophone, or a 10 dB decrease in the gain of the hydrophone array or the receiver will all have an identical effect on the detection function. The consequence of this 10 dB change is to reduce the distance over which whales are certain to be detected from 0.7 to 0.3 km and to reduce the maximum distance at which whales can be detected from 4 to 3.1 km. Similarly, an increase in the SNR by 10 dB would increase the lower limit of the detection function from 0.7 to 1.3 km, and the upper limit from 4.0 to 4.9 km. The shape of the detection function between these two end points remains nearly unchanged.

The most important component in the sonar equation is the on-axis source level of the clicking whale for which a nominal value of $200 \text{ dB}_{\text{rms}} \text{ re } 1 \mu\text{Pa}$ at 1 m was assumed in constructing Fig. 3. However, the source level of regular clicks during deep dives is not constant. For example, the source level of sperm whale echolocation clicks is known to vary by up to 20 dB (Madsen *et al.*, 2002) and the apparent output of beaked whale echolocation clicks seems to vary by a similar factor (Madsen *et al.*, 2005). However, the effect of SL on SNR is strictly additive. Therefore, the consequence of a statistical variation in SL will be to decrease the distance at which whales are certain to be detected and to increase the maximum detection distance.

Altering the RT has a similar effect to changing SL, but on the probability of detection axis of Fig. 3. Increasing RT will decrease the distance at which whales can be detected with certainty (possibly to zero) and decrease the maximum distance at which whales can be detected but will also reduce the rate of false alarms.

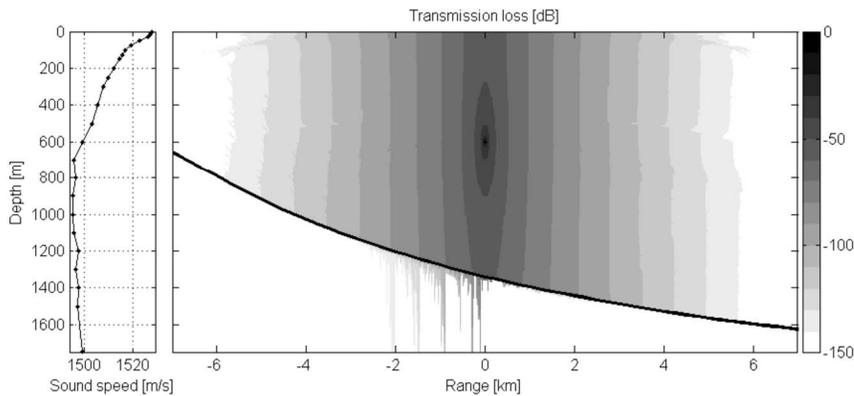


FIG. 4. TL estimation using an average sound-speed profile (left panel). The TL (right panel) was estimated using Bellhop (Porter and Bucker, 1987) for a 40 kHz omnidirectional signal with an assumed bottom profile as shown.

The ambient noise level, which masks the whale sound during reception, is the second most important parameter in the sonar equation. Ambient noise is very difficult to model. Most textbook descriptions of sonar performance base their high-frequency noise level estimation on the Knudsen model for surface-generated noise (Urlick, 1983; Burdic, 1984; Lurton, 2002). Actual noise levels are very often higher than predicted by the Knudsen model due to additional noise sources such as rain, thunderstorms, and shipping (Wenz curves; Urlick, 1983; Aguilar *et al.*, 2006). It is essential to measure the ambient noise level *in situ* at the frequencies of interest to predict its impact on the sonar performance in any particular site.

Surface-generated noise decreases as a function of depth. At the frequencies used by beaked whales for echolocation, surface-generated ambient noise is reduced by more than 20 dB for a hydrophone depth of 2000 m [Lurton, 2002, Eq. (4.41)]. This means that deep hydrophones, in general, will perform better than shallow ones provided they are no more than 1 km below the whale foraging depth. Deep hydrophones will also be more likely to detect the downward-directed clicks produced during descent. However, operational constraints (see below) may limit the depth at which hydrophones can be deployed. The impact of ambient noise may also be reduced by increasing the gain of the array (AG) and/or the receiver (PG). Increasing AG requires increased investment in hardware (more hydrophones, more processing power, etc.), thus reducing the cost effectiveness of PAD.

PG is related to the analysis bandwidth, which is typically adapted to the type of signal to be detected. The maximum PG can be estimated based on the signal characteristics of beaked whale clicks. Both species analyzed so far (Cuvier's and Blainville's beaked whales) show a unique time-variable spectrum of short frequency-modulated upsweeps (Zimmer *et al.*, 2005a; Johnson *et al.*, 2006), at least when recorded close to the acoustic axis. To take advantage of this feature, a PAD processor could implement a matched filter to emphasize beaked whale signals resulting in a PG of $10 \log(BT)$, where BT is the time-bandwidth product. For Cuvier's beaked whale, this would be about 9 dB (signal bandwidth of 40 kHz and signal length of 0.2 ms). This gain may be reduced when echolocation clicks are detected at large off-axis angles because frequency dependent distortions would then become significant. As a result, the performance of the

matched filter would decrease due to increasing mismatch between the received signal and its replica in the filter.

In our calculations, TL was modeled using the spherical spreading law. Strictly speaking, this is only valid in an unbounded medium with constant sound speed. Such theoretical conditions never exist in real oceans, especially for long distances, where the presence of ocean boundaries, refraction, and scattering results in the need for complicated models of TL. Nevertheless, the spherical spreading law is a reasonable first approach in assessing the performance of PAD over the relatively short distances in question here. However, it should be recognized that performance at longer ranges may increase due to sound trapping or decrease due to shadow zones (e.g., see Lurton, 2002, Fig. 2.25 for propagation in deep sound channel and Fig. 2.27 for shadow and convergence zones). Prediction of TL for any specific site requires acoustic propagation modeling that includes realistic sound-speed profiles in combination with a detailed description of the bottom characteristics of the site. Figure 4 shows as example a TL estimation for a late summer sound-speed profile typically found in subtropical Atlantic climate (28° latitude). The Bellhop model (Porter and Bucker, 1987) predicts for a beaked whale foraging at 600 m depth, a TL of about 130 dB at a horizontal distance of 5 km.

The final relevant component of the sonar equation is the attenuation resulting from off-axis transmission of echolocation clicks. The sonar of echolocating whales, measured to date, is characterized by rather narrow sound beams with directivity indices normally more than 25 dB (Au, 1983; Zimmer *et al.*, 2005a, 2005b). This is equivalent to a half-power beam width of less than 10° (Zimmer *et al.*, 2005a) and results in a substantial decrease in apparent source level at large off-axis angles. For Cuvier's beaked whale, the maximum off-axis attenuation has been measured to be about 46 dB (Zimmer *et al.*, 2005a). This narrow beam width translates to a low probability that a random single click would be detected at a distance by a single PAD system, as shown clearly in Fig. 2. Fortunately, echolocating whales emit a large number of clicks in varying directions during each foraging dive. This substantially increases the probability that a given PAD system will detect at least one click, as shown by our modeling of the consequence of the variations in orientation during the foraging portion of a dive (Fig. 3).

The depth at which beaked whales forage (d), and therefore produce clicks consistently, is likely to vary both within and, particularly, between dives, and from location to location. Dive depth will affect both the slant range to the hydrophone of the PAD system [Eq. (12)] and the elevation of the hydrophone from the whale [η , Eq. (9)]. However, these effects are greatest at distances of less than 1 km (i.e., in the region where detection of at least one click is certain). At distances of greater than 1 km, the effect of a $\pm 10\%$ variation in d is relatively small ($\pm 1^\circ$ for η and $\pm 10\%$ for R) provided the difference between whale and hydrophone depth is not too great. We therefore argue that the shape of the detection function is not particularly sensitive to variations in whale depth.

B. Detection functions

The maximal detection range (4 km in Figs. 2 and 3) and the minimal range, within which no echolocation clicks are missed (0.7 km), depend only on the two extremes of the off-axis attenuation and are independent of the actual distribution of off-axis angles. The maximal range where whales may be detected corresponds to the limit for the detection of on-axis clicks, and the guaranteed detection range, where echolocation clicks are always detectable, results from clicks observed with the maximum off-axis attenuation.

The actual shape of the detection function depends on the precise form of the PDF for off-axis angles. Here we have modeled this distribution for a single complete dive, containing 4000 clicks, so that the resulting detection function indicates the probability of detecting at least one click during the dive. The resulting detection function indicates that about 25% of dives produce clicks that are close enough to on axis to be detected at a range of up to 3.8 km. A further 25% of the dives produced clicks so far off axis that they can only be detected when the animal is within 0.8 km of the hydrophone.

We compared the simulated detection function (solid black line in Fig. 3) with a detection function derived from measured values for the depth, pitch, and heading for each click made by real Cuvier's beaked whales during the course of foraging dives (gray lines in Fig. 3). When we examined the distribution of headings for these animals, we found that it did not correspond very closely to that used to generate the simulation (solid black line in Fig. 3). In particular, some of the whales tended to reverse heading by about 180° , as might be expected if they were carrying out an area-restricted search through a prey concentration. This behavior resulted in a PDF of heading that had two modes. The circular normal distribution does not account for such inversions of direction. We therefore developed another kind of simulated dive during which 50% of the clicks had a mean heading γ_i and the remaining 50% had a mean heading ($\gamma_i + 180^\circ$). The resulting detection function is shown by the dashed black curve in Fig. 3.

The two simulated detection functions (with and without reversal of swimming direction (solid and dashed lines, respectively, in Fig. 3) bracket the empirical detection func-

tions derived from direct measurements of beaked whale dives, suggesting that our method provides a realistic model for estimating general detection functions.

Beaked whales initiate regular clicking during the descent phase of a deep dive, at a depth of about 400–500 m, usually well above the foraging layer (Tyack *et al.*, 2006) whereby these clicks are usually downward oriented and are therefore likely to be detected only by hydrophones below the mean foraging depth or at very short ranges by hydrophones located close to the surface, as indicated for large off-axis angles in Fig. 2.

C. Classification of detected clicks

Any real detector produces false alarms and in practical applications, it is necessary to evaluate each detection to assess the likelihood that it represents a real signal. Detecting beaked whales then involves a two-step process: detection of clicklike sounds and classification of these sounds as stemming from a beaked whale. Classification performance can be enhanced not only by considering the spectral features of individual clicks, as described above, but also by analyzing the temporal characteristics of click sequences, taking advantage of the observation that echolocating animals tend to produce sequences of clicks with a slowly varying interclick interval (ICI).

Temporal classification requires the detection of consecutive clicks to estimate ICI. As Fig. 1 indicates, the acoustic beam produced by beaked whales is rather narrow (here modeled with a 6° beamwidth at -3 dB). As a result, a receiver is likely to only record a series of clicks (known as a “scan”) over a short time interval. Clicks within scans can vary in level by more than 20 dB as the beam crosses the hydrophone location (Johnson *et al.* 2006, Madsen and Wahlberg, 2007). The objective of detecting at least one click made by an echolocating whale can easily be transformed to one of detecting at least one scan if RT is increased by 20–30 dB. Different odontocete species have differing but overlapping distributions of ICIs. Cuvier's beaked whales have a mean ICI of 0.4 s (Zimmer *et al.*, 2005a) whereas sperm whales tend to use ICIs > 0.5 s (Zimmer *et al.*, 2005b). Dolphins typically have an ICI of less than 0.1 s (Madsen *et al.*, 2004) and so might be relatively easily differentiated from beaked whales on this basis.

Single click classification is, in general, spectrally based and requires a good SNR over an extended signal band to be successful. The spectral description of Cuvier's beaked whale click (Zimmer *et al.*, 2005a) indicates that spectral energy 10 dB below the peak signal value may be sufficient to differentiate a single click from those of other nonziphiid echolocating toothed whales. Increasing RT by 10 dB would ensure that only signals that allow spectral classification are detected.

An advanced detector should be able to adapt to the environment and change its parameter settings (e.g., detection threshold and analysis bandwidth) according to the success or failure of past decisions. Wald (1947) described a decision scheme based on a sequential probability ratio test, which does not require every detection opportunity (e.g., ev-

ery detected transient) to be classified. If it is uncertain, the detector may defer a decision until after the reception of a new data set. This scheme is repeated until the desired decision (beaked whale present or absent) can be made. Such a scheme would initially try to classify the detection according to spectral features but, if uncertain, would change its classification strategy as the number of clicks with the “correct” ICI in a sequence increases.

The relatively low detection threshold ($RT=14$ dB) used here will give a high false alarm rate, making classification essential for most applications. However, the addition of a classifier will modify the overall detection function and it is important to re-evaluate the detection function for the entire PAD system once the detection-classification algorithms are established.

As discussed above, the performance of a PAD hydrophone should improve as its depth is increased because of the lower ambient noise. This suggests that hydrophones mounted on, or close to, the bottom, such as those at the AUTECH naval underwater range (Moretti *et al.*, 2006), should outperform hydrophones close to the surface. However, operational considerations may favor shallower hydrophone deployments. For example, it is difficult to deploy a towed hydrophone at or below the mean foraging depth of beaked whales. An additional advantage of a hydrophone that is situated above the whale is that it may be able to exploit surface reflections of upward-directed clicks to provide an additional classification cue. The delay between the direct arrival of a click and its surface reflection depends only on the geometry between whale and the hydrophone, which should remain fairly constant during the short duration of a click sequence. The resulting stable double click could therefore be a useful cue for classification of clicks stemming from an individual deep-diving whale. The decreased SNR of a shallow hydrophone may therefore be compensated by an improved classification performance, which would allow a reduction in RT.

D. Complicating factors

The above analyses assume that clicks are emitted by a single animal. It is therefore important to understand what will happen when more than one animal is present. A naïve suggestion would be that detection probability scales directly with the number of clicking animals. However, the level of improvement will depend on the extent to which the animals' orientations are correlated and on the performance of the receiver in the presence of multiple animals.

Surface observations indicate that members of beaked whale groups tend to surface and dive in close vicinity to each other but little is known about subsurface behavior of groups of whales, especially with respect to their interindividual relative orientation. Zimmer *et al.* (2005a) showed that a pair of Cuvier's beaked whales remained within a few hundred meters of each other during a foraging dive. As a result, the cumulative acoustic footprint of a group, i.e., the acoustically illuminated volume, and hence the probability that it will be detected, may change little with group size.

The presence of multiple foraging animals may also

confound the classification process, especially when classification is based on ICI. However, the length of time that a single whale directs detectable clicks toward the hydrophone is likely to be limited to a few seconds at a time due to the narrow sonar beam and the constant movement of the animals. Consequently, provided that individual whales orient independently, their clicks will be received by a hydrophone at different times, thus reducing the interference. Within-scan interference by conspecifics can be eliminated by increasing the RT so that only high-level on-axis clicks are used for classification purposes. Nonetheless, the potential difficulties associated with classification based on temporal cues underline the need for classification concepts that are based on spectral or joint information.

In addition to conspecific interference, there is also the potential of interference from other cetacean species. For example, dolphins also emit short echolocation clicks that are characterized by a very broadband spectrum that, in most cases, overlaps with the spectrum of clicks emitted by beaked whales. However, discrimination of dolphins from beaked whales may be aided by the presence of whistles, their shorter ICIs, and the lack of significant click energy below 20 kHz in beaked whale clicks (Zimmer *et al.*, 2005a, Johnson *et al.*, 2006).

E. Operational considerations

The way in which a PAD system is operated will influence detection performance. A key operational parameter is how long the PAD system should monitor a water volume before making a decision about the presence or absence of the target species. This depends on the specific purpose of the whale detector. If the main aim is the mitigation of anthropogenic impacts on deep-diving whales, then an obvious sampling interval is one that is long enough to include at least one foraging dive and the interval between such dives.

Foraging beaked whales echolocate in regular bouts of duration about 30 min for Cuvier's beaked whales (Tyack *et al.*, 2006), but the time between consecutive deep dives, that is the time where whales are silent, is variable and can exceed 110 min. This yields an initial requirement for a listening time of at least 140 min in order to have a high probability of detecting Cuvier's beaked whales. The optimal listening time may be longer both to sample multiple dive cycles and because our estimates of silent durations are drawn from a limited data set and therefore may be biased toward short intervals.

If PAD is to be used as part of a survey whose purpose is to estimate whale density, then it is desirable to estimate the distance of the whale from the hydrophone when it is detected. This can be done either by detecting the same whale simultaneously with multiple PAD systems and determining its location by triangulation or by exploiting information from the multipath structure of sound propagation, if this is available. For example, a compact three-dimensional array of hydrophones may estimate the direction to the whale and can make use of a single surface reflection to estimate the range and depth of a clicking whale.

All of the calculations presented here have been based on the assumption that the PAD hydrophone is stationary. If this assumption is relaxed, the assessment of PAD performance becomes more complicated and requires detailed modeling of the relative motion of the whale and hydrophone to determine a dynamic acoustic detection function. The acceptable platform speed will depend on the goals of PAD. For risk mitigation requiring a high level of confidence that the absence of detections was indeed due to the absence of whales, relatively low PAD platform speeds, probably less than 4 km h^{-1} , are needed to ensure adequate time in each water volume. These speeds may be best achieved by mounting hydrophones on deep-diving gliders or drifting buoys. On the other hand, if PAD is being used to determine whale density, it may be more cost effective to tow a hydrophone behind a vessel. The reduced probability of detecting whales when they are present would be compensated by the ability to survey a larger area within a given time or for a given cost.

IV. CONCLUSIONS

At present, PAD offers the only reliable method for determining the presence or absence of deep-diving echolocating beaked whales, which are difficult, and often impossible, to detect visually. Here we show that the performance of a PAD system can be characterized using a detection function, which describes how the probability of detecting an acoustically active whale varies as function of its distance from the receiver, derived using the sonar equation. This indicates that a beaked whale is almost certain to be detected during a foraging dive if it is within approximately 0.7 km of a PAD system and unlikely to be detected if it is more than 4 km away for the assumed propagation conditions and click source properties. The precise limits for detection depend on the source level of the whale's clicks, the ambient sound level at the depth of the hydrophone, and the gain settings of the hydrophone arrays and receiver. Between these limits, the form of the detection function depends on the way in which the orientation of the whale varies with respect to the PAD hydrophone during the course of a dive. We show, by comparison with detection functions derived using data collected directly from diving whales, that these variations can be represented by a simple statistical model.

Some of the factors that affect the performance of PAD (the quality and quantity of the listening equipment) are under the control of the observer. However, the whales control their sound emissions and the environment determines the propagation and the background noise level, which masks the whale's echolocation clicks. Because the performance of PAD depends heavily on environmental conditions, it must be implemented carefully to compensate, at least partially, for the effect of these conditions.

The availability of beaked whales for acoustic detection depends on the interval between foraging dives. This silent interval may exceed 110 min so that extended listening periods will be required to ensure a high probability that whales are detected by a PAD system. This suggests that if a PAD system is being used to mitigate the risks to beaked

whales associated with anthropogenic sound, it will be most effective if its hydrophones are mounted on slow moving platforms (such as gliders or drifting buoys) rather than a fast moving active sonar vessel.

The detection functions presented here are based on a single choice of parameters. Empirical detection functions could and should be used to calibrate the different model assumptions. As more dive data are measured from beaked whales, statistical analysis of relevant dive parameters may be used to improve the associated PDF. As different applications (risk mitigation or abundance estimation) have different requirements, it will also be important to analyze the sensitivity of their success to the actual shape of the detection function.

ACKNOWLEDGMENTS

This work was carried out within the Marine Mammal Risk Mitigation Project of the NATO Undersea Research Center. We thank Len Thomas and Tiago Marques for comments and suggestions. Cuvier's beaked whale data were collected by the Woods Hole Oceanographic Institution and BluWest S.R.L. with funding from SERDP and the National Oceanographic Partnership Program (NOPP). P.T.M. was funded by a Steno grant from the National Danish Science Foundation. M.P.J. and P.L.T. were funded under the Office of Naval Research, SERDP and the NOPP. Tagging was performed under US National Marine Fisheries Service research Permit Nos. 981-1578-02 and 981-1707-00 to P.L.T. and with the approval of the Woods Hole Oceanographic Institution Animal Care and Use Committee.

¹An off-axis angle $\vartheta=0$ is "on axis."

²But see [Urick \(1962\)](#) for a general sonar equation suited also for transients; see also [Au \(1993\)](#).

- Aguilar de Soto, N., Johnson, M., Madsen, P. T., Tyack, P. L., Bocconcelli, A., and Borsani, J. F. (2006). "Does intense ship noise disrupt foraging in deep-diving Cuvier's beaked whales (*Ziphius cavirostris*)?" *Marine Mammal Sci.* **22**, 690–699.
- Au, W. W. (1993). *The Sonar of Dolphins* (Springer, New York).
- Barlow, J. (1999). "Trackline detection probability for long-diving whales," in *Marine Mammal Survey and Assessment Methods*, edited by G. W. Garner, S. C. Amstrup, J. L. Laake, B. F. J. Manly, L. L. McDonald, and D. G. Robertson (A.A. Balkema, Rotterdam), pp. 209–224.
- 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. L. (2005). "Estimates of sperm whale abundance in the northeastern temperate Pacific from combined acoustic and visual survey," *Marine Mammal Sci.* **21**, 429–445.
- 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, UK).
- Burdic, W. S. (1984). *Underwater Acoustic System Analysis* (Prentice-Hall, Englewood Cliffs, NJ).
- Cox, T. M., Ragen, T. J., Read, A. J., Vos, E., Baird, R. W., Balcomb, K., Barlow, J., Caldwell, J., Cranford, T., Crum, L., D'Amico, A., D'Spain, G., Fernández, A., Finneran, J., Gentry, R., Gerth, W., Gulland, F., Hildebrand, J., Houser, D., Hullar, T., Jepson, P. D., Ketten, D., MacLeod, C. D., Miller, P., Moore, S., Mountain, D. C., Palka, D., Ponganis, P., Rommel, S., Rowles, T., Taylor, B., Tyack, P., Wartzok, D., Gisiner, R., Mead, J., and Benner, L. (2006). "Understanding the impacts of anthropogenic sound on beaked whales," *J. Cetacean Res. Manage.* **7**, 177–187.
- Drouot, V., Gannier, A., and Goold, J. C. (2004). "Summer social distribu-

- tion of sperm whales (*Physeter macrocephalus*) in the Mediterranean Sea." *J. Mar. Biol. Assoc. U.K.* **84**, 675–680.
- Fisher, N. I. (1993). *Statistical Analysis of Circular Data* (Cambridge University Press, Cambridge, UK).
- Gordon, J. C. D. (1987). "The behaviour and ecology of sperm whales off Sri Lanka," Ph.D. thesis, Darwin College, Cambridge.
- Johnson, M., Madsen, P. T., Zimmer, W. M. X., Aguilar de Soto, N., 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., Aguilar de Soto, N., 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.
- 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.
- Lucifredi, I., and Stein, P. J. (2007). "Gray whale target strength measurements and the analysis of the backscattered response," *J. Acoust. Soc. Am.* **121**, 1383–1391.
- Lurton, X. (2002). *An Introduction to Underwater Acoustics* (Springer, London).
- Madsen, P. T., Johnson, M., Aguilar de Soto, N., Zimmer, W. M. X., and Tyack, P. L. (2005). "Biosonar performance of foraging beaked whales (*Mesoplodon densirostris*)," *J. Exp. Biol.* **208**, 181–194.
- Madsen, P. T., Kerr, I., and Payne, R. (2004). "Echolocation clicks of two free-ranging, oceanic delphinids with different food preferences: False killer whales (*Pseudorca crasidens*) and Risso's dolphins (*Grampus griseus*)," *J. Exp. Biol.* **207**, 1811–1823.
- Madsen, P. T., and Wahlberg, M. (2007). "Recording and quantitative analysis of clicks from echolocating toothed whales in the wild," *Deep-Sea Res., Part II* **54**, 1421–1444.
- Madsen, P. T., Wahlberg, M., and Møhl, B. (2002). "Male sperm whale (*Physeter macrocephalus*) acoustics in a high-latitude habitat: Implications for echolocation and communication," *Behav. Ecol. Sociobiol.* **53**, 31–41.
- Mellinger, D. K., Stafford, K. M., and Fox, C. G. (2004). "Seasonal occurrence of sperm whale (*Physeter macrocephalus*) sounds in the gulf of Alaska, 1999–2001," *Marine Mammal Sci.* **20**, 48–62.
- Møhl, B., Wahlberg, M., Madsen, P. T., Miller, L. A., and Surlykke, A. (2000). "Sperm whale clicks: Directionality and source level revisited," *J. Acoust. Soc. Am.* **107**, 638–648.
- Moretti, D., Morrissey, R., DiMarzio, N., and Ward, J. (2006). "Verified passive acoustic detection of beaked whales (*Mesoplodon densirostris*) using distributed hydrophones in the tongue of the ocean, Bahamas (A)," *J. Acoust. Soc. Am.* **119**, 3374.
- Porter, M. B., and Bucker, H. P. (1987). "Gaussian beam tracing for computing ocean acoustic fields," *J. Acoust. Soc. Am.* **82**, 1349–1359.
- Teloni, V. (2005). "Patterns of sound production in diving sperm whales in the northwestern Mediterranean," *Marine Mammal Sci.* **21**, 446–457.
- Thode, A. (2004). "Tracking sperm whale (*Physeter macrocephalus*) dive profiles using a towed passive acoustic array," *J. Acoust. Soc. Am.* **116**, 245–253.
- Tyack, P. L., Johnson, M., Soto, N. A., Sturlese, A., and Madsen, P. T. (2006). "Extreme diving of beaked whales," *J. Exp. Biol.* **209**, 4238–4253.
- Urick, R. J. (1962). "Generalized form of the sonar equations," *J. Acoust. Soc. Am.* **34**, 547–550.
- Urick, R. J. (1983). *Principles of Underwater Sound*, 3rd ed. (McGraw-Hill, New York).
- Wahlberg, M. (2002). "The acoustic behaviour of diving sperm whales observed with a hydrophone array," *J. Exp. Mar. Biol. Ecol.* **281**, 53–62.
- Wald, A. (1947). *Sequential Analysis* (Wiley, New York).
- Watkins, W. (1980). "Acoustics and the behavior of sperm whales," in *Animal Sonar Systems*, edited by R. Busnel and J. F. Fish (Plenum, New York), pp. 291–297.
- Zimmer, W. M. X., Johnson, M. P., D'Amico, A., and Tyack, P. L. (2003). "Combining data from a multisensor tag and passive sonar to determine the diving behavior of a sperm whale (*Physeter macrocephalus*)," *IEEE J. Ocean. Eng.* **28**, 13–28.
- Zimmer, W. M. X., Johnson, M. P., Madsen, P. T., and Tyack, P. L. (2005a). "Echolocation clicks of free-ranging Cuvier's beaked whales (*Ziphius cavirostris*)," *J. Acoust. Soc. Am.* **117**, 3919–3927.
- Zimmer, W. M. X., Tyack, P. L., Johnson, M. P., and Madsen, P. T. (2005b). "Three-dimensional beam pattern of regular sperm whale clicks confirms bent-horn hypothesis," *J. Acoust. Soc. Am.* **117**, 1473–1485.