# Sperm whale three-dimensional track, swim orientation, beam pattern, and click levels observed on bottom-mounted hydrophones[a)]

Eva-Marie Nosal[b)] and L. Neil Frazer
*School of Ocean and Earth Science and Technology, University of Hawaii, Honolulu, Hawaii 96822*

(Received 2 May 2007; revised 22 July 2007; accepted 2 August 2007)

In an earlier paper [Nosal and Frazer Appl. Acoust. **61**, 1187–1201 (2006)], a sperm whale was tracked in three-dimensions using direct and surface-reflected time differences (DRTD) of clicks recorded on five bottom-mounted hydrophones, a passive method that is robust to timing errors between hydrophones. This paper refines the DRTD method and combines it with a time of (direct) arrival method to improve the accuracy of the track. The position and origin time of each click having been estimated, pitch and yaw are then obtained by assuming the main axis of the whale is tangent to the track. Roll is then found by applying the bent horn model of sperm whale phonation, in which each click is composed of two pulses, p0 and p1, that exit the whale at different points. With instantaneous pitch, roll, and yaw estimated from time differences, amplitudes are then used to estimate the beam patterns of the p0 and p1 pulses. The resulting beam patterns independently confirm those obtained by Zimmer *et al.* [J. Acoust. Soc. Am. **117**, 1473–1485 (2005); **118**, 3337–3345 (2005)] with a very different experimental setup. A method for estimating relative click levels is presented and used to find that click levels decrease toward the end of a click series, prior to the "creak" associated with prey capture.
© *2007 Acoustical Society of America.* [DOI: 10.1121/1.2775423]

PACS number(s): 43.30.Sf, 43.60.Jn, 43.80.Ka [WWA] Pages: 1969–1978

## I. INTRODUCTION

The main purpose of this paper is to demonstrate the progress and use of passive acoustic methods for studying marine mammals in the wild, especially odontocetes. In a recent paper (Nosal and Frazer, 2006), we studied the improvement in ray-based tracking that occurs when a realistic sound speed profile is used instead of an assumed isospeed profile. We tracked a sperm whale using the difference between direct and surface-reflected click arrival times (DRTD), a method that is robust to time-origin errors on different hydrophones. Here we refine the DRTD method and combine it with a time of (direct) arrival (TOA) method to get a combined method that is more accurate than either method separately. The time and position estimates are precise enough that we can approximate swim velocity and orientation at each click, which we then use to estimate click beam patterns and levels.

This paper focuses on clicks with regular interclick intervals of 0.45–2 s, called "usual" clicks by Whitehead and Weilgart (1990), emitted by sperm whales while foraging at depth. A typical foraging dive lasts about 45 min, and begins with a steep, steady descent to a depth of several hundred meters (Watkins *et al.*, 2002; Watwood *et al.*, 2006), followed by a period of searching for prey and then a steep, steady ascent. Series of regular clicks emitted at foraging depth are often terminated by a "creak" of clicks with high repetition rate followed by several seconds of silence (Gordon, 1987; Mullins *et al.*, 1988; Goold and Jones, 1995). The regular clicks are likely used for echolocation (Norris and Harvey, 1972; Whitehead and Weilgart, 1991; Goold and Jones, 1995; Møhl *et al.*, 2000; Jaquet *et al.*, 2001; Madsen *et al.*, 2002a; Whitehead, 2003; Møhl *et al.*, 2003) while the creaks mark the terminal phase of prey capture (Miller *et al.*, 2004). Regular clicks are short in duration, broadband (100 Hz to over 20 kHz), and have a powerful forward directed beam (Møhl *et al.*, 2000; Madsen *et al.*, 2002a, b). Regular clicks are often heard on hydrophones several kilometers from the vocalizing animal. Being of short duration, direct and reflected arrivals can often be distinguished, making clicks ideal for passive localization.

We track a single sperm whale from its regular clicks for 23 min using recordings on five bottom-mounted hydrophones. The data were recorded at the Atlantic Undersea Test and Evaluation Center (AUTEC) located in the Tongue of the Ocean (off Andros Island, Bahamas). They were provided by the Naval Undersea Warfare Center for the second International Workshop of Detection and Localization of Marine Mammals using Passive Acoustics. The sampling rate was 48 kHz and the hydrophone positions are listed in Table I. Some further details can be found in Adam *et al.* (2006), but unfortunately, the anti-alias filter, frequency response, sensitivity, and directionality of the sensors were not available. Accordingly, our results are limited by the assumption of an omnidirectional and flat frequency response, and absolute sound pressure levels cannot be found.

[a)]Portions of this work were presented in "Bottom-mounted hydrophones used to investigate sperm whale click and swim characteristics," Proceedings of the Fourth Joint Meeting of the Acoustical Society of America and the Acoustical Society of Japan, Honolulu, HI, Nov. 2006.

[b)]Author to whom correspondence should be addressed; electronic mail: nosal@hawaii.edu

TABLE I. Hydrophone positions provided by NUWC.

| Hydrophone | $x$ position (m) | $y$ position (m) | Depth (m) |
|---|---|---|---|
| G | 10 658.04 | −14 953.63 | 1530.55 |
| H | 12 788.99 | −11 897.12 | 1556.14 |
| I | 14 318.86 | −16 189.18 | 1553.58 |
| J | 8 672.59 | −18 064.35 | 1361.93 |
| K | 12 007.50 | −19 238.87 | 1522.54 |

The data set used here is the same data set that we used to develop the DRTD method (Nosal and Frazer, 2006). The track of the sperm whale in this data set has also been obtained using time of arrival differences between pairs of receivers (Giraudet and Glotin, 2006; Morrissey *et al.*, 2006; White *et al.*, 2006). The improved accuracy of the combined DRTD/TOA method used in this paper allows us to estimate the velocity of the sperm whale from position and time differences between successive clicks. The pitch and yaw of the whale then follow from the assumption that the main axis of the whale is parallel to its velocity vector.

In order to find roll, we apply the bent-horn model of sperm whale phonation, which was proposed to explain the multipulse structure of sperm whale clicks (Norris and Harvey, 1972; Møhl, 2001) and is supported by recent studies (Madsen *et al.*, 2003; Møhl *et al.*, 2003; Zimmer *et al.*, 2005b). In the bent-horn model, a single sound is generated at the phonic lips (Fig. 1). Some energy leaks directly into the water as a p0 pulse. Most of the energy transmits back through the spermaceti organ, reflects off the frontal sac in front of the skull, transmits forward into the "junk," and exits into the water as the p1 pulse. Since the p1 pulse follows a longer path than p0, it arrives later, giving the click a multipulse structure. Other click components (resulting from further reflections in the head and other exit points) are also present (Møhl, 2001), but they are not required or considered here. Since the measured delay between the p0 and p1 pulses depends on the orientation of the whale (Zimmer *et al.*, 2005a; Laplanche *et al.*, 2006), it can be used to recover roll.

With position, velocity, pitch, roll and yaw obtained solely from travel time differences, we then use relative amplitudes to estimate the beam patterns and directivity indices of the p0 and p1 pulses that comprise a click. Our results agree with previous studies (Møhl *et al.*, 2003; Zimmer *et al.*, 2005b) which found that the p1 pulse has a narrowly focused, forward-direct beam, that the p0 pulse is slightly weaker and more broadly backward-directed, and that a low-frequency, nearly omnidirectional component is characteristic of all clicks. Finally, we correct click amplitudes for beam pattern and propagation loss to estimate relative click levels within each click sequence, finding that click source levels decrease toward the end of a click series.

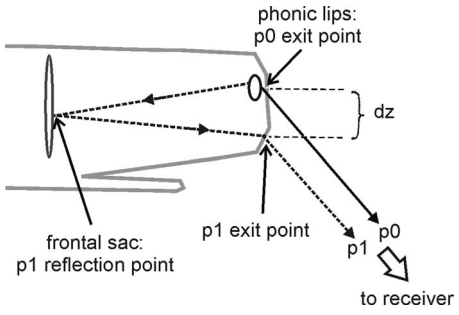



FIG. 1. Diagram illustrating the paths taken by the p0 and p1 pulses according to the modified bent-horn model of sound production in sperm whales.

## II. DATA PROCESSING

### A. Click detection, classification, and association

The beginnings and ends of the clicks were detected using an automated transient detector (Page, 1954; Wald, 1947; Abraham, 2000; Zimmer *et al.*, 2003, 2005b). To reduce noise, each time series was filtered using a second-order, high-pass Butterworth filter with a 300 Hz low cut. The envelope of each filtered time series was calculated as the magnitude of the corresponding analytic signal, where the analytic signal has real and imaginary parts consisting of the original time series and its Hilbert transform, respectively (e.g., Bracewell, 2000). Given the instantaneous signal amplitude (envelope) $e_n$, a test variable $V_n$ was calculated as

$$V_n = \frac{e_n^2}{N_n}, \tag{1}$$

where $N_n$ is the noise estimate. For the first noise estimate, $N_1$ is taken as the mean square envelope value (over all samples for that hydrophone). Detection decisions and updates for subsequent noise estimates are made according to the value of $V_n$ in the following algorithm

$$\text{If } \begin{cases} V_n > T_0 \\ V_n < T_x \\ \text{other} \end{cases} \text{ then } \begin{cases} \text{decide detection} \\ \text{set } T_x = T_1 \\ \text{set } N_{n+1} = N_n \\ \text{decide no detection} \\ \text{set } T_x = T_2 \\ \text{set } N_{n+1} = (1-\alpha)N_n + \alpha e_n^2 \\ \text{no decision} \\ \text{keep current } T_x \\ \text{set } N_{n+1} = N_n \end{cases} \tag{2}$$

where $T_0$, $T_1$, and $T_2$ are the thresholds for decision of detection, end of detection, and noise, respectively; $T_x \in \{T_1, T_2\}$ is the current threshold; and $\alpha$ is the exponential weighting on the power estimate when no signal is detected. For the first sample, $T_x$ is set equal to $T_1$. Threshold and weighting values that performed well were $T_0 = 25(13.98\ \text{dB})$, $T_1 = 9(9.54\ \text{dB})$, $T_2 = 4(6.02\ \text{dB})$, and $\alpha = 1/100$.

At each time step, this algorithm decides if there is a signal present (detection) or not (no detection). The algorithm operates in two modes: signal and noise. In the signal mode, signal present is decided while the value of the test variable is greater than the detection threshold ($V_n > T_0$). No decision is made if the value of the test variable is less than the threshold for detection but greater than the threshold for

the end of detection ($T_1 < V_n < T_0$). Signal not present is decided once the value of the test variable drops below the end of detection threshold ($V_n < T_1$). Here, the algorithm switches to noise mode. In this mode, signal not present is decided while the test variable remains below the noise threshold ($V_n < T_2$) and the noise variable is updated at each time step according to an exponential weighting (more weight toward recent values). No decision is made if the value of the test variable is less than the threshold for detection but greater than the threshold for noise ($T_2 < V_n < T_0$). Signal present is decided once the value of the test variable jumps above the detection threshold ($V_n > T_0$). Here, the algorithm switches to signal mode.

Each click resulted in a direct arrival, usually followed by a lower-amplitude surface reflection. Direct-reflected pairs were classified according to the following criteria: (1) the amplitude of the direct arrival varies slowly; (2) the interclick interval between successive direct arrivals varies slowly; (3) the reflected arrival has lower amplitude than the direct arrival; and (4) the time between the direct and reflected arrival (DRTD) is similar to that of the preceding direct-reflected pair.

Each click detected on two or more receivers was numbered sequentially. Clicks on different receivers were associated by comparing intervals between clicks, which should be nearly identical on all receivers. To eliminate incorrect associations due to click time measurement error, this comparison included intervals between all clicks in a series, not only those immediately preceding or following a given click. In total, 1324 clicks were numbered, with 1102, 913, 868, 1163, and 1035 clicks detected on receivers 1, 2, 3, 4, and 5, respectively. The number of clicks detected on a total of two, three, four, and five receivers was 137, 324, 480, and 383. Only clicks recorded on three or more hydrophones (total of 1187 clicks) were used for localization.

### B. Level and pulse delay measurement

The maximum of the envelope was used to estimate the received peak pressures of each direct click. Using frequency and time windows, peak pressures of the p0, p1, and low-frequency (LF) components were also obtained. Following Zimmer *et al.* (2005b), the p0 and p1 pulses were defined to fall in time windows from, respectively, −2 to 3 ms and 3 to 10 ms, relative to detection of the start of the click. They were both defined to fall in a frequency window of 3–22 kHz. The identified p0 and p1 times and amplitudes corresponded to maxima of the envelope of the filtered signal. The LF component was defined with a time window of −2 to 10 ms and a frequency window of 300 Hz to 3 kHz. The delay between the p0 and p1 pulses, $\tau$, was estimated by subtracting p0 arrival time from p1 arrival time.

## III. LOCALIZATION

### A. Methods

For each receiver, the acoustic propagation model BELLHOP (Porter, 2005) was used to create a lookup table of TOAs, DRTDs, takeoff beam angles, and transmission losses for a list of candidate source ranges and depths. The historic depth-dependent sound speed profile for the area (24° 45′ N, 77° 45′ W) in March was taken from the Generalized Digital Environment Model (GDEM) and is the same as the profile used in Nosal and Frazer (2006). The depth list varied from 5 to 1550 m with 5 m increments, and the range list varied from 5 m to 20 km with 5 m increments. Since arrival times varied smoothly for the depths and ranges of interest, all required TOAs and DRTDs were interpolated from the values in the lookup table using cubic splines.

To determine the time and position of each click, we first created a four-dimensional grid of candidate source points (one dimension for time and three for position). Errors in DRTD and TOA were assumed to be normally distributed. Ideally, DRTD and TOA should be regarded as functions, not just of source position, but also of sound speed and receiver positions, and likelihood surfaces should be maximized over this much larger parameter space. However, to reduce computational requirements, we incorporated the uncertainties in sound speed profile and receiver positions into the standard deviations for DRTD and TOA in a "worst case" manner. Standard deviations $\sigma_{\mathrm{DRTD}}$ and $\sigma_{\mathrm{TOA}}$ were calculated as

$$\sigma_{\mathrm{DRTD}} = \sqrt{2\sigma_{\mathrm{meas}}^2 + \sigma_{\mathrm{rp}}^2 + \sigma_{\overline{\mathrm{DRTD}}}^2}, \tag{3}$$

$$\sigma_{\mathrm{TOA}} = \sqrt{\sigma_{\mathrm{meas}}^2 + \sigma_{\mathrm{rp}}^2 + \sigma_{\overline{\mathrm{TOA}}}^2}, \tag{4}$$

where $\sigma_{\mathrm{meas}}$ is the standard deviation (s.d.) in the measured click times, $\sigma_{\mathrm{rp}}$ is the s.d. due to uncertainty in receiver position, and $\sigma_{\overline{\mathrm{DRTD}}}$ and $\sigma_{\overline{\mathrm{TOA}}}$ are the s.d. (due to sound speed uncertainty) in modeled DRTDs and TOAs. We used $\sigma_{\mathrm{meas}} = 5$ ms based on the widths of the clicks (about 10 ms) and $\sigma_{\mathrm{rp}} = 2$ ms corresponding approximately to a best-guess receiver position uncertainty of 3 m (actual position uncertainty is unknown). To determine $\sigma_{\overline{\mathrm{DRTD}}}$ and $\sigma_{\overline{\mathrm{TOA}}}$, the DRTD and TOA lookup tables were recalculated for all 12 months using historic sound speed profiles (also from the GDEM). This gave 12 possible TOAs and DRTDs for each range and depth. The difference between the minimum and the maximum of these 12 values approximates the width of the uncertainty curves. The maximum such width over all ranges and depths ("worst-case") was taken as 2 s.d., giving $\sigma_{\overline{\mathrm{DRTD}}} = \sigma_{\overline{\mathrm{TOA}}} = 3$ ms. Using the maximum width simplifies the calculations, by allowing 1 s.d. to be used for all candidate points, and it overestimates final errors.

For candidate whale position $\mathbf{s}$ and click time $t$, the DRTD and TOA likelihood functions were computed as

$$L_{\mathrm{DRTD}}(\mathbf{s}) = \frac{1}{(2\pi\sigma_{\mathrm{DRTD}}^2)^{N/2}} \times \exp\left[\frac{-1}{2\sigma_{\mathrm{DRTD}}^2} \sum_j (\overline{\mathrm{DRTD}}_j(\mathbf{s}) - \mathrm{DRTD}_j)^2\right], \tag{5}$$

$$L_{\mathrm{TOA}}(\mathbf{s},t) = \frac{1}{(2\pi\sigma_{\mathrm{TOA}}^2)^{N/2}} \times \exp\left[\frac{-1}{2\sigma_{\mathrm{TOA}}^2}\sum_j (\overline{\mathrm{TOA}}_j(\mathbf{s},t) - \underline{\mathrm{TOA}}_j)^2\right], \tag{6}$$

where the sums are over all receivers that heard the click, $N$ is the number of receivers that heard the click, $\underline{\mathrm{DRTD}}_j$ and $\underline{\mathrm{TOA}}_j$ are the measured values on receiver $j$, and $\overline{\mathrm{DRTD}}_j$ and $\overline{\mathrm{TOA}}_j$ are the modeled values. The total likelihood value is the product of these:

$$L(\mathbf{s},t) = L_{\mathrm{DRTD}}(\mathbf{s})L_{\mathrm{TOA}}(\mathbf{s},t). \tag{7}$$

The point $(\mathbf{s},t)$ that maximizes $L$ is the estimated source position and time. An advantage of distinct likelihood surfaces is that they can be examined separately as a diagnostic, since persistent differences between locations from the two methods are an indication that hydrophone time origins may be different (degrading TOA), or that the sound speed profile in the upper part of the water column is inaccurate (degrading DRTD).

For computational efficiency, two passes were made. The first pass was coarsely sampled in space (10 m grid spacing) and time (10 ms time spacing). For the first click, the spatial search volume covered the full water column in depth and extended 5 km past the boundary defined by the receivers. Time was searched from 0 to 20 s. For the other clicks the boundary of the search volume was based on the time, $\Delta t$, between the current click and the preceding localized click. This was estimated from the measured time between these two clicks on a single phone that heard both clicks. The search volume was centered on the position estimate of the previous click, and bounded in all three directions by double the maximum possible swim distance in $\Delta t$, i.e., $8\Delta t$ (assuming a swim speed of at most 4 m/s). Time was searched from the previously localized click until $2\Delta t$ after it.

The second pass refined the position and time estimate from the first pass by searching a smaller, more finely sampled, volume centered on the position and time found in the first pass. The search volume for this pass was sampled at intervals of 1 m in space and 1 ms in time. It was bounded in space by the coarsely determined source location, plus or minus 20 m in all directions, and in time by 200 ms before and after the coarsely determined click time.

### B. Error estimates

The literature on bioacoustic localization arrays contains various approaches to quantify error (e.g., Whalberg *et al.*, 2001; Spiesberger and Wahlberg, 2002). Since the complete likelihood surfaces were already calculated in the above-noted localization step, we applied a somewhat different approach, using the likelihood surfaces to give error estimates.

95% confidence intervals (CIs) were estimated from conditional likelihood functions (CLFs) by identifying the interval containing 2.5%–97.5% of the cumulative likelihood for the parameter of interest. For example, to find the confidence interval in the $x$ position for a single click, all other parameters ($y$ position, $z$ position, and click time) were fixed to their values $(y_0,z_0,t_0)$ at the estimated source position and time. The corresponding CLF, $L(x_j|y_0,z_0,t_0)$, was calculated according to Eq. (7) for a list of possible $x$ positions, $x_j$. The cumulative CLF was then calculated as

$$C(x) \approx \frac{\sum_{x_j<x} L(x_j)}{\sum_{x_j} L(x_j)}. \tag{8}$$

The denominator normalizes the distribution, and the equality is approximate because of the discrete sampling of $x$. The list of $x$ positions ranged from $x_0-1$ km to $x_0+1$ km (since the CLF was very close to 0 at 1 km away from $x_0$), with 1 m resolution. Then the 95% CI is $[x_{2.5\%},x_{97.5\%}]$ where $x_{2.5\%}$ and $x_{97.5\%}$ are such that $\mathrm{C}(x_{2.5\%})=0.025$ and $\mathrm{C}(x_{97.5\%})=0.975$. CIs for $y$ position, $z$ position, and time were computed similarly.

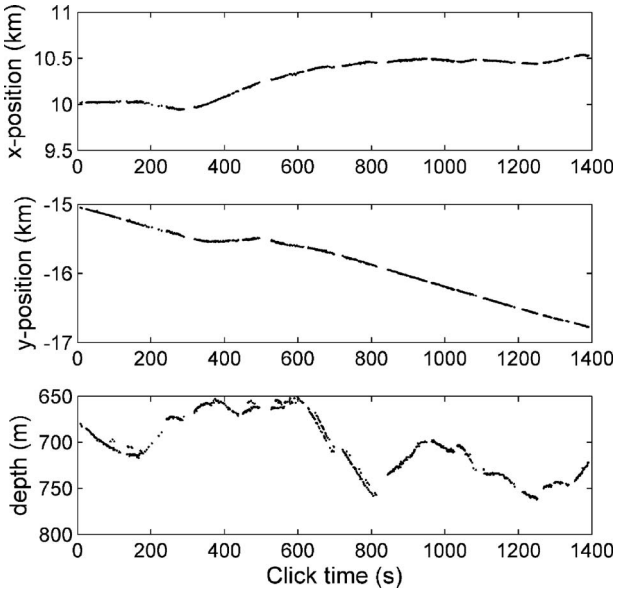



FIG. 2. Estimated track (not smoothed) of the sperm whale. Positions of clicks detected on three or more receivers are plotted as dots against the time of the click.

### C. Results

Figure 2 shows the resulting $x$, $y$, and $z$ positions obtained for clicks heard on three or more receivers. Figure 3 shows the positions in the $x$-$y$ plane. The click time list ranged from $t_0-1$ s to $t_0+1$ s with 0.1 ms resolution. The resulting CI half-widths for position are shown in Fig. 4. The half-widths for time were less than 4.5, 5, and 5.8 ms for clicks heard on five, four, and three receivers, respectively.

## IV. SWIM ORIENTATION

### A. Conventions

This section outlines the orientation conventions used in the following. Two reference frames are required: the earth frame and the whale frame, notated as unprimed and primed, respectively. In the earth frame, positive $x$, $y$, and $z$ are di-

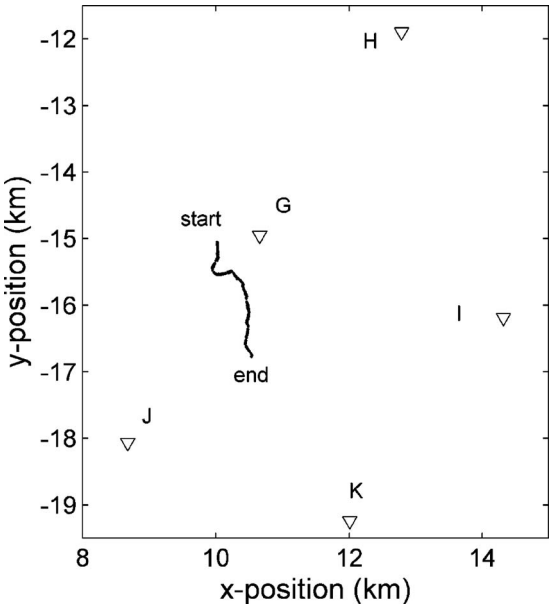



FIG. 3. Estimated two-dimensional track (not smoothed) of the sperm whale. Positions are plotted as dots to form the track, as in Fig. 2, and triangles indicate receiver positions.

rected east, north, and upward, respectively. In the whale frame, positive $x'$, $y'$, and $z'$ point forward (rostrally along the whale's long axis), left, and dorsally, respectively. The two frames coincide when the whale is traveling due east, horizontal and upright.

Three angles are required to transform between the whale and earth frames: yaw ($\theta_z$), pitch ($\theta_y$), and roll ($\theta_x$), which are rotations about the $z$, $y$, and $x$ axes, respectively. For yaw and roll, positive values correspond to a coordinate system rotation in a clockwise direction when looking away from the origin along the axis of rotation. For consistency with conventions used by Johnson and Tyack (2003), and so that a positive pitch corresponds to a nose-upward orientation, positive pitch corresponds to a counter-clockwise rotation when looking away from the origin along the $y$ axis. Note that this convention differs from standard Euler and pitch-roll-heading convention (Goldstein, 1980). Thus in our convention a whale with zero yaw, pitch, and roll is swimming eastward, horizontally, and upright. From this $\theta_x=\theta_y=\theta_z=0$ orientation, the whale turns left to increase yaw, toward the surface to increase pitch, and clockwise to increase roll. To make them unique, $\theta_x$, $\theta_y$, and $\theta_z$ are constrained to the intervals $(-180°, 180°]$, $[-90°, 90°]$, $(-180°, 180°]$, respectively.

A vector in the earth (unprimed) frame is expressed in whale (primed) frame coordinates via three matrices that commute only in the limit of very small angles (so the order of multiplication is important):

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = R_x(\theta_x)R_y(\theta_y)R_z(\theta_z)\begin{bmatrix} x \\ y \\ z \end{bmatrix} \tag{9}$$

with

$$\mathrm{R}_x(\theta_x) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_x & \sin\theta_x \\ 0 & -\sin\theta_x & \cos\theta_x \end{bmatrix}, \tag{10}$$

$$\mathrm{R}_y(\theta_y) = \begin{bmatrix} \cos\theta_y & 0 & \sin\theta_y \\ 0 & 1 & 0 \\ -\sin\theta_y & 0 & \cos\theta_y \end{bmatrix}, \tag{11}$$

$$\mathrm{R}_z(\theta_z) = \begin{bmatrix} \cos\theta_z & \sin\theta_z & 0 \\ -\sin\theta_z & \cos\theta_z & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{12}$$

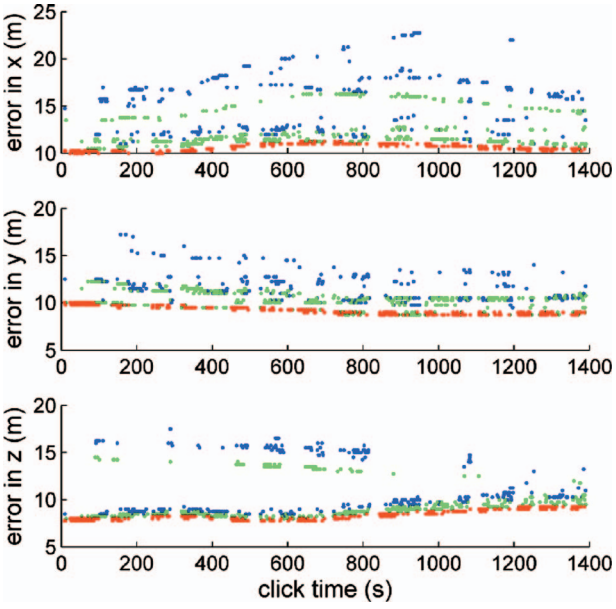



FIG. 4. Half-widths of 95% confidence intervals for position. Red, green, and blue indicate clicks detected on five, four, and three receivers, respectively.

### B. Pitch and yaw

The first step in recovering swim attitude is to approximate the velocity of the whale at each click. To do this, a vector-valued position function $\mathbf{f}(t)=(f_x(t), f_y(t), f_z(t))$ was fit to the calculated click positions and times by minimizing a weighted sum of squared position error and acceleration:

$$E(\mathbf{f}) = a\sum_{j=1}^{N_c} |\mathbf{s}_j - \mathbf{f}(t_j)|^2 + (1-a)\int_{t_1}^{t_{N_c}} \left|\frac{d^2\mathbf{f}}{dt^2}\right|^2 dt, \tag{13}$$

where the sum is over all localized clicks $j$, $N_c$ is the total number of localized clicks, $\mathbf{s}_j$ and $t_j$ are the estimated position and time of click $j$, and $a=0.7$ is a smoothing parameter. Velocity in the earth frame, $\mathbf{v}(t)=(v_x(t), v_y(t), v_z(t))$, is found by taking the first derivative of $\mathbf{f}$.

To recover pitch and yaw, we assume that the whale's main axis is parallel to its velocity. This assumption neglects the effects of current and the ability of the whale to move laterally or vertically, as well as any scanning movements of the head, so the goodness of the approximation increases with the forward speed of the whale. Pitch and yaw can then be computed as

$$\theta_y = \sin^{-1}(v_z/|\mathbf{v}|), \tag{14}$$

$$\theta_z = \tan^{-1}(v_y/v_x), \tag{15}$$

where $\tan^{-1}$ is the four-quadrant, inverse tangent.

### C. Roll

Once position, pitch, and yaw are known, roll is estimated from the delay $\tau$ between the p0 and p1 pulses (Fig. 1). To do this, we build on methods introduced by Zimmer *et al.* (2005a) and Laplanche *et al.* (2006). The modeled delay is split into two components: $\tau = t_c + t_{\Delta p}$. The constant component, $t_c$, is the time required for sound to travel from the phonic lips to the frontal sac, where it reflects, and thence forward to the p1 exit point. It is assumed fixed for a given animal. The second component, $t_{\Delta p}$, is the difference between the travel time from the p1 exit point to the receiver and the travel time from the p0 exit point to the receiver. It depends on the exit point of the p1 pulse relative to the phonic lips—we assume this exit point is fixed for a given animal—and on whale position, receiver position, and roll. The exit point of the p1 pulse is located at the junk and is directly ventral to the phonic lips (Madsen, 2002; Zimmer *et al.*, 2005a). Hence, if we take the phonic lips to be at point (0,0,0) in the whale frame, the p1 exit point can be approximated as $(0,0,-dz)$, with $dz$ fixed for the individual whale (Fig. 1).

Given the position of the whale, **s**, and the receiver, **r**, in the earth frame, as well as the pitch and yaw angles previously determined, $\tau$ can be modeled (for various $t_c$, $dz$, and roll angles) as follows. For each click and receiver, we find the takeoff direction of the ray that connects the receiver and localized source positions. For a constant sound-speed profile, this direction in the earth frame would simply be $\mathbf{r}-\mathbf{s}$ with source position **s** and receiver position **r**. To find this direction for our depth-dependent sound speed profile, the list of takeoff angles from BELLHOP is interpolated to get the vertical angle (or elevation), $\phi$, of the ray in the earth frame. The ray direction vector in the earth frame is then

$$\mathbf{b} = [r_x - s_x,\ r_y - s_y,\ \tan(\phi)\sqrt{(r_x - s_x)^2 + (r_y - s_y)^2}]^T. \tag{16}$$

This ray direction is transformed into the whale frame direction vector, $\mathbf{b}'$, by applying Eq. (9) with the calculated yaw, pitch, and candidate roll angle for the current click. The azimuth $\alpha' \in [-180°, 180°]$ and elevation $\phi' \in [-90°, 90°]$ of the ray in whale coordinates are then calculated as

$$\alpha' = \tan^{-1}\left(\frac{b'_y}{b'_x}\right), \quad \phi' = \tan^{-1}\left(\frac{b'_z}{\sqrt{(b'_x)^2 + (b'_y)^2}}\right), \tag{17}$$

where $\tan^{-1}$ is the four-quadrant inverse tangent. Here, positive/negative azimuth corresponds to a leftward/rightward directed beam, and positive/negative elevation corresponds to an upward/downward directed beam. Elevation and azimuth of 0° correspond to a beam directed along the whale's main axis, $x'$.

Since the distance from the whale to the receiver is much greater than $dz$, the vertical takeoff angles of the p0 pulse and the p1 pulse from the junk exit point are well approximated by $\phi'$. Then $t_{\Delta p}$ is approximated as (Fig. 5)

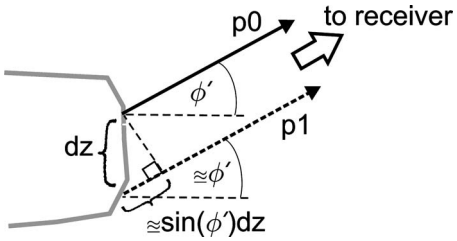



FIG. 5. Angle approximation used to estimate the delay between p0 and p1.

$$t_{\Delta p} \cong \sin(\phi')dz/c \tag{18}$$

in which $c$ is the speed of sound through water ($c_w$), for $-90° \leqslant \alpha' \leqslant 90°$, but $c$ is the variable speed of sound through whale tissue ($c_t$), for $\alpha' > 90°$ or $\alpha' < -90°$. The change in sound speed is necessary because for clicks propagating forward ($-90° \leqslant \alpha' \leqslant 90°$) both the p0 path, and the p1 path after exiting the junk, pass primarily through the water, while backward propagating pulses pass through whale tissue. For each click, $c_w$ is found from the value of the sound speed profile interpolated to the depth of the whale, while $c_t$ is a function of temperature, pressure, ray elevation, and ray azimuth (different angles mean that sound passes through different tissues). For simplicity, the unknown value of $c_t$ is assumed here to be constant and is estimated in the following optimization step.

The constants $t_c$, $dz$, and $c_t$, and the roll for each click are found as follows. With $t_c$, $dz$, and $c_t$ fixed over all clicks, we find the roll for each click that minimizes the difference between measured and modeled $\tau$ in a least-squares sense over all receivers. Summing over all clicks gives the total squared error associated with the current values of $t_c$ and $dz$. This total error is minimized over $t_c$, $dz$, and $c_t$.

The best fit values were $t_c = 6.6$ ms, $dz = 1.30$ m, and $c_t = 1540$ m/s. The value $t_c = 6.6$ ms corresponds to a whale length of 14.42 m using the formulas of Gordon (1991), and to a whale length of 13.61 m using the formula of Rhinelander and Dawson (2004). The estimated p1 exit point located 1.30 m ventral of the phonic lips for a whale over 13.5 m makes sense anatomically assuming that the exit point is on the junk (Møhl, 2001). It is also consistent with the results of Zimmer *et al.* (2005a), who found the p1 exit point to be 1.10 m ventral of the phonic lips for a 12 m whale. The derived value $c_t = 1540$ m/s is high compared to the value of 1370 m/s found by Flewellen and Morris (1978) for the speed of sound through spermaceti oil at 33 °C at 1 atm. It is more consistent with (although still on the high end of) values for more similar conditions given by Goold *et al.* (1996), who found that sound speed in spermaceti oil increased from 1390 to 1540 m/s with increasing pressure (from 0 to 90 atm) and decreasing temperature (from 38 to 22 °C). The seemingly high value for $c_t$ found here possibly stems from the fact that our animal is alive, and that the p1 pulse passes through other whale tissue (not only spermaceti oil) to get to the receiver; however, we have made numerous assumptions and approximations that invariably introduce error, and our estimate will need to be examined in

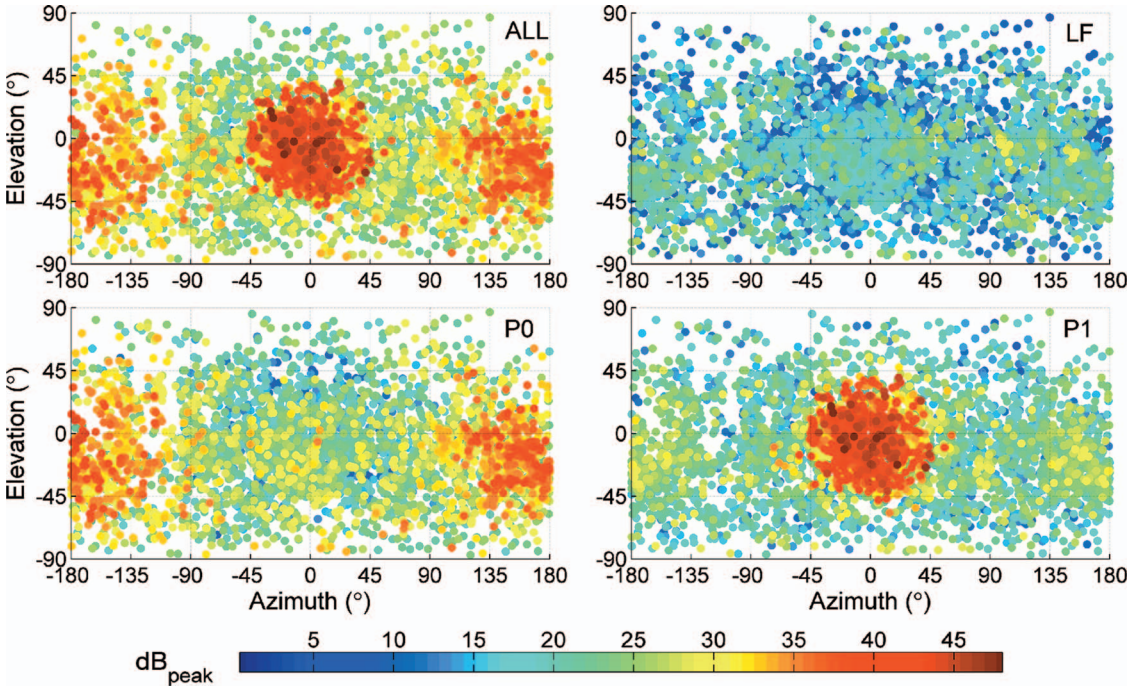



FIG. 6. Estimated beam patterns from 4807 recorded clicks. Since receiver sensitivities were not available, these are not absolute click levels but relative levels, such that 0 dB corresponds to the weakest recorded click. Recorded levels were corrected for transmission loss and were plotted as colored dots on the appropriate elevation/azimuth position. Higher levels are shown overlapping lower levels to minimize the effect of a variable source level. Results are shown for the full click, p0 pulse, p1 pulse, and LF components.

future work.

## V. BEAM PATTERN AND DIRECTIVITY

### A. Methods and results

The azimuth and elevation of each click to each receiver were found for the calculated position and orientation data as outlined in Sec. IV. The received levels obtained in Sec. II B were corrected for transmission loss using the values in the lookup table from Sec. III A to get click levels. Since hydrophone sensitivity (or clipping level) was unavailable, click levels could only be found as values relative to some arbitrary level, chosen such that the weakest click level corresponded to 0 dB. Hence, we report only "relative click levels," by which we mean the difference between the current click level and the minimum click level (over all clicks).

Relative click levels are plotted as a function of azimuth and elevation in Fig. 6. Since 324, 480, and 383 clicks were localized on three, four, and five receivers, respectively, a total of $324\times3+480\times4+383\times5=4807$ points are plotted. Multiple clicks with similar azimuth and elevation were measured, and the figures show higher levels overlapping lower levels, which helps to reduce the effect of variation in click source levels by approximating the maximum level in each direction. The resulting beam patterns are similar, although with somewhat broader peaks, to the patterns found by Zimmer *et al.* (2005b), who used a similar approach.

Although absolute source levels cannot be estimated, source level differences can be found: The maximum source level of the p1 pulse measured here was 8.8 dB higher than the maximum source level of the p0 pulse and 19.4 dB higher than the maximum source level of the LF component. These values are consistent with estimates reported by Zimmer *et al.* (2005b) of 210 $dB_{peak}$ for the p1 pulse, 200 $dB_{peak}$ for the p0 pulse, and 190 $dB_{peak}$ for the LF component, all re: 1 $\mu$Pa at 1 m.

Directivity indices were estimated according to a discretized version of Eq. (3-10) of Au (1993)

$$\mathrm{DI} = 10\log\frac{4\pi}{\sum_{i=1}^{N_\alpha}\sum_{j=1}^{N_\phi}\left(\frac{p(\alpha_i',\phi_j')}{p_{\max}}\right)^2\cos\alpha'\Delta\alpha'\Delta\theta'},$$

in which $N_\alpha$ is the number of azimuth steps of width $\Delta\alpha'$, $N_\phi$ is the number of elevation steps of width $\Delta\phi'$, $p(\alpha_i',\phi_j')$ is the received pressure corrected for transmission loss for the bin corresponding to azimuth step $\alpha_i'$ and elevation step $\phi_j'$ (recall that the primes denote whale frame coordinates), and $p_{\max}$ is the maximum received pressure over all angles. Step widths of 2.5° were used for both azimuth and elevation, and the maximum pressure over each bin was used for $p(\alpha_i',\phi_j')$. Estimated directivity indices were 21.8 dB for the p1 pulse, 9.4 dB for the p0 pulse, and 5.2 dB for the LF component. In comparison, Møhl *et al.* (2003) reported a p1 directivity index of 27 dB and Zimmer *et al.* (2005b) reported a p1 directivity index of 26.7 dB and a p0 directivity index of 7.4 dB.

### B. Discussion

Similar to the case discussed in Zimmer *et al.* (2005b), it is likely that the maximum source level of the p1 pulse is underestimated here due to clipping of the high-intensity arrivals (197 out of all 4807 signals used reached clipping amplitude), limited sampling bandwidth, and a small sample size of on-axis clicks (only 8 clicks within 5° of the main

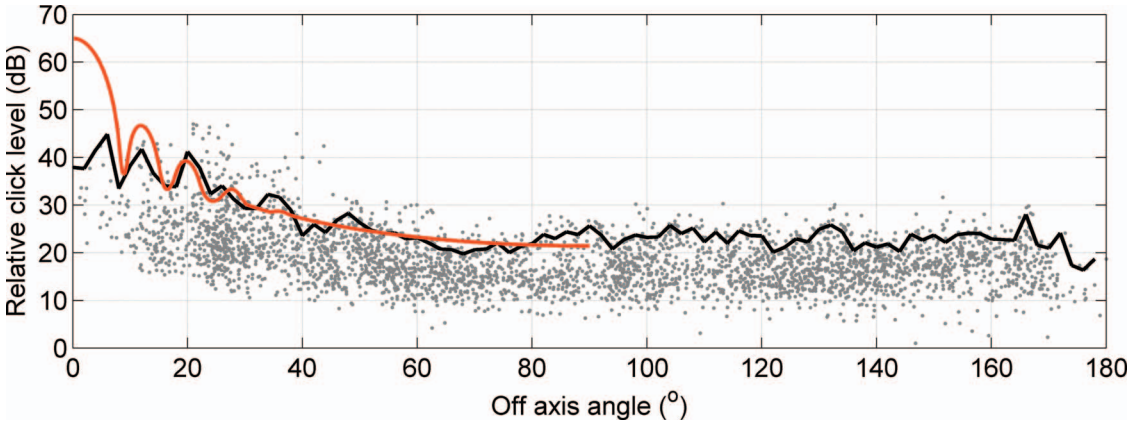



FIG. 7. Scatter plot of estimated p1 beam pattern as a function of off-axis angle. Levels are relative, as in Fig. 6. The black line represents the 90th percentile for each off-axis angle bin (bin size 2°). The red line represents the beam pattern predicted for a circular piston with parameters fitted to measured values (black line) for off-axis angles between 20° and 90°. Results and best-fit parameters are given in the text.

beam axis). This explains in part why our beam pattern for the p1 pulse has a broader peak and lower directivity than reported by Zimmer *et al.* (2005b) and Møhl *et al.* (2003).

To deal with clipping, we follow Zimmer *et al.* (2005b), in which a model broadband beam pattern is fit to the 90th percentile for off-axis angles between 20° and 90° (binned here into 2° intervals). In Fig. 7, which is similar to Fig. 9 of Zimmer *et al.* (2005b), all measured (relative) levels are plotted with gray dots as a function of off-axis angle. The 90th percentile for each (2°) off-axis bin is plotted in black and the best-fit modeled beam pattern is plotted in red. The p1 pulse was modeled as a Gabor function emitted from a circular piston (Au, 1993). The parameters that give the best fit (in a least-squares sense) are peak frequency 15 kHz, signal duration 0.60 ms, and piston radius 0.40 m. These correspond to a p1 directivity index of 25.2 dB. This is much closer to, although still considerably less than, the values of 27 and 26.7 dB reported by Møhl *et al.* (2003) and Zimmer *et al.* (2005b), respectively. Unfortunately, maximum source level could not be estimated here because hydrophone sensitivity was unknown.

Other sources of error in our beam pattern are from errors in estimated pitch, yaw, and roll. These stem from uncertainties in source location and click time, from the assumption that the whale's main axis coincides with the velocity vector, and from approximations made in application of the bent-horn model (constant sound speed through whale tissue, for example). Further, our results are limited by the assumption that all clicks have the same beam pattern and by the assumptions on receiver response (flat-frequency, omnidirectional, and the same for all receivers). The remarkable agreement of our beam patterns with those of Zimmer *et al.* (2005b) and Møhl *et al.* (2003), both of which used very different experimental setups, suggests that either our assumptions are at least approximately satisfied or that the errors caused by the assumptions tend to cancel over multiple hydrophones.

## VI. CLICK SOURCE LEVELS

Click source levels were estimated from the measured beam patterns by finding best-fit levels to a "model" beam pattern. For a given direction, the model beam pattern was assigned the maximum received level, corrected for transmission loss, over all directions within 5°. This approach was preferred over binning the received levels into discretized azimuth and elevation steps, which would have resulted in a nonuniform weighting of the received levels since different elevation bins subtend different solid angles (this also explains why the clicks in Fig. 6 are more densely populated at elevations closer to 0°). For a sufficient number of clicks, this should eliminate variations due to click source level, giving a model that well approximates the true beam pattern (Zimmer *et al.* 2005b).

Relative click source levels (at a distance of 1 m on the acoustic axis) were estimated by minimizing the misfit between the model levels and the received levels corrected for transmission loss and source level. Minimization was done in a least-squares sense over all receivers that heard the click. Again, because receiver sensitivity was unknown, only relative click source levels could be found; the resulting relative click source levels are shown in Fig. 8. A total of 14 complete click series and 2 incomplete click series (at the beginning and the end of the data set) were recorded, where a series is defined as ending in a creak or at least 5 s of silence. Click levels vary by about 20 dB [in agreement with the dynamic range reported by Madsen *et al.* (2002)] and tend to steadily decrease toward the end of each click series. There were no apparent correlations between the interclick intervals and source levels, or between whale depth (or orientation) and source level. However, as shown in Fig. 9, there is a significant relationship between click level and the order of

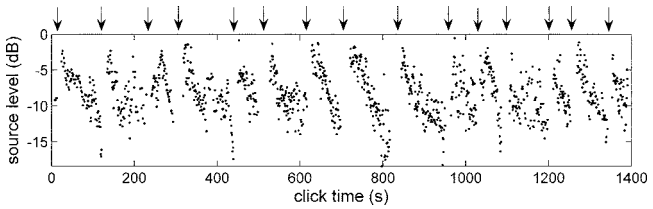



FIG. 8. Source levels relative to the strongest recorded source level as a function of time. Click levels decline by 10–15 dB from the start to the end of most click series. The beginning of each series is indicated by an arrow at the top.

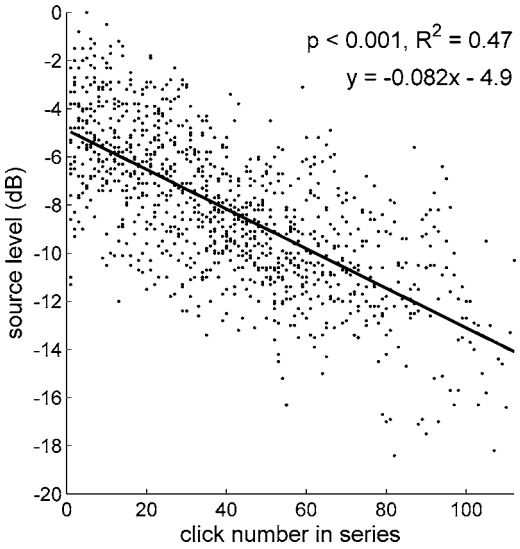



FIG. 9. Relative source level as a function of click number within its series. Data are pooled from all 14 complete click series. The significant correlation and the negative slope of the regression line suggest that click level decreases with click number within a series.

the click within its series. This suggests that the variation in click level may be a consequence of the click production mechanism, whereby a click series begins at some constant level and decreases with each subsequent click. However, other explanations are possible. For example, since we do not have target range information it is not possible to determine if level is controlled by some automatic gain control mechanism, as might be employed by dolphins (Au and Benoit-Bird, 2003; Au and Herzing, 2003) but possibly not by beaked whales (Madsen *et al.*, 2005).

Recalculating the directivity indices after correcting with these click source levels gave indices of 22.9, 9, and 5 dB for the p1, p0, and LF components, respectively. Since corrected beam patterns were very similar to those in Fig. 6 they are not presented.

## VII. CONCLUDING REMARKS

Although our method to recover roll is specific to sperm whales, our estimation of pitch and yaw is applicable to any clicking marine mammal. Since beam patterns often exhibit rotational symmetry, at least to a first approximation—as for sperm whales (Zimmer *et al.* 2005b; this paper) and for bottlenose dolphins (Au, 1993)—it may be useful to estimate beam patterns as a function of off-axis angle only. In that case, roll is not needed, so the methods developed here can be used to obtain directivity indices and estimates of click level for any clicking marine mammal recorded on bottom-mounted hydrophones in the wild, provided that the clicks are heard on enough hydrophones to produce 3D locations.

## ACKNOWLEDGMENTS

This work was funded by the U.S. Office of Naval Research. Data were provided by the Naval Undersea Warfare Center. The authors thank two anonymous reviewers for comments that helped improve the clarity of the final manuscript.

Abraham, D. A. (**2000**). "Passive acoustic detection of marine mammals," SM-351, SACLANT Undersea Research Centre, La Spezia, Italy.

Adam, O., Motsch, J.-F., Desharnais, F., DiMarzio, N., Gillespie, D., and Gisiner, R. C. (**2006**). "Overview of the 205 workshop on detection and localization of marine mammals using passive acoustics," Appl. Acoust. **67**, 1061–1070.

Au, W. W. L. (**1993**). *The Sonar of Dolphins* (Springer, New York.

Au, W. W. L., and Benoit-Bird, K. (**2003**). "Automatic gain control in the echolocation system of dolphins," Nature (London) **423**, 861–863.

Au, W. W. L., and Herzing, D. L. (**2003**). "Echolocation signals of wild Atlantic spotted dophin (*Stenella frontalis*)," J. Acoust. Soc. Am. **113**, 598–604.

Bracewell, R. N., (**2000**). "The Hilbert transform," in *The Fourier Transform and Its Applications*, 3rd ed. (McGraw-Hill, Singapore), pp. 359–364.

Flewellen, C. G., and Morris, R. J. (**1978**). "Sound velocity measurements on samples from the spermaceti organ of the sperm whale *Physeter catodon*," Deep-Sea Res. **25**, 269–277.

Giraudet, P., and Glotin, H. (**2006**). "Real-time 3D tracking of whales by precise and echo-robust TDOAs of clicks extracted from 5 bottom-mounted hydrophones records of the AUTEC," Appl. Acoust. **67**, 1106–1117.

GDEM (Generalized Digital Environment Model) website (last accessed on 1 May, 2007), ⟨https://128.160.23.42/gdemv/gdemv. html⟩.

Goldstein, H. (**1980**). "The Euler angles" and "Euler angles in alternate conventions," in *Classical Mechanics*, 2nd ed. (Addison-Wesley, Reading, MA), Sec. 4–4, pp. 143–148 and Appendix B, pp 606–610.

Goold, J. C., Bennell, J. D., and Jones, S. E. (**1996**). "Sound velocity measurements in spermaceti oil under the combined influences of temperature and pressure," Deep-Sea Res., Part I **43**, 961–969.

Goold, J. C., and Jones, S. E. (**1995**). "Time and frequency domain characteristics of sperm whale clicks," J. Acoust. Soc. Am. **98**, 1279–1291.

Gordon, J. C. (**1987**). "The behaviour and ecology of sperm whales off Sri Lanka," Ph.D. thesis, University of Cambridge, Cambridge, UK.

Gordon, J. C. D. (**1991**). "Evaluation of a method for determining the length of sperm whales (*Physeter catodon*) from their vocalizations," J. Zool. **224**, 301–314.

Jaquet, N., Dawson, S., and Douglas, L. (**2001**). "Vocal behavior of male sperm whales: Why do they click?," J. Acoust. Soc. Am. **109**, 2254–2259.

Johnson, M., 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.

Laplanche, C., Adam, O., Lopatka, M., and Motsch, J.-F. (**2006**). "Measuring the off-axis angle and the rotational movements of phonating sperm whales using a single hydrophone," J. Acoust. Soc. Am. **119**, 4074–4082.

Madsen, P. T. (**2002**). in "Sperm whale sound production," Ph.D. thesis, University of Aarhus, Denmark.

Madsen, P. T., Carder, D. A., Au, W. W. L., Nachtigall, P. E., Møhl, B., and Ridgway, S. H. (**2003**). "Sound production in neonate sperm whale," J. Acoust. Soc. Am. **113**, 2988–2991.

Madsen, P. T., Johnson, M., Aguilar deSoto, N., Zimmer, W. M. X., and Tyack, P. (**2005**). "Biosonar performance of foraging beaked whales (*Mesoplodon densirostris*)," J. Exp. Biol. **208**, 181–194.

Madsen, P. T., Payne, R., Kristiansen, N. U., Wahlberg, M., Kerr, I., and Møhl, B. (**2002a**). "Sperm whale sound production studied with ultrasound time/depth-recording tags," J. Exp. Biol. **205**, 1899–1906.

Madsen, P. T., Wahlberg, M., and Møhl, B. (**2002b**). "Male sperm whale (*Physeter macrocephalus*) acoustics in a high-latitude habitat: Implications for echolocation and communication," Behav. Ecol. Sociobiol. **53**, 31–41.

Miller, P. J. O., Johnson, M. P., and Tyack, P. L. (**2004**). "Sperm whale behaviour indicates the use of echolocation click buzzes 'creaks' in prey capture," Proc. R. Soc. London, Ser. B **271**, 2239–2247.

Møhl, B. (**2001**). "Sound transmission in the nose of the sperm whale *Physeter catodon*. A postmortem study," J. Comp. Physiol. [A] **187**, 335–340.

Møhl, B., Wahlberg, M., Madsen, P. T., Heerfordt, A., and Lund, A. (**2000**). "Sperm whale clicks: Directionality and source level revisited," J. Acoust. Soc. Am. **114**, 1143–1154.

Møhl, B., Wahlberg, M., Madsen, P. T., Heerfordt, A., and Lund, A. (**2003**). "The monopulsed nature of sperm whale clicks," J. Acoust. Soc. Am. **114**, 1143–1154.

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.

Mullins, J., Whitehead, H., and Weilgart, L. (**1988**). "Behaviour and vocalizations of two sperm whales, Physeter macrocephalus, off Nova Scotia," Can. J. Fish. Aquat. Sci. **45**, 1736–1743.

Norris, K. S., and Harvey, G. W. (1972). "A theory for the function of the spermaceti organ of the sperm whale (*Physeter catodon L.*)," in *Animal Orientation and Navigation*, edited by S. R. Galler, K. Schmidt-Koenig, G. J. Jacobs, and R. E. Belleville, NASA Special Publication No. **262**, 397–417.

Nosal, E.-M. and Frazer, L. N. (**2006**). "Track of a sperm whale from delays between direct and surface-reflected clicks," Appl. Acoust. **67**, 1187–1201.

Page, S. E. (**1954**). "Continuous inspection schemes," Biometrika **41**, 100–115.

Porter, M. (**2005**). "BELLHOP" (a Gaussian beam/finite element beam code). Available in the *Acoustics Toolbox* at http://www.hlsresearch.com/oalib/Modes/AcousticsToolbox/, Last accessed 9/18/2007.

Rhinelander, M. Q., and Dawson, S. M. (**2004**). "Measuring sperm whales from their clicks: Stability of interpulse intervals and validation that they indicate whale length," J. Acoust. Soc. Am. **115**, 1826–1831.

Spiesberger, J. L., and Wahlberg, M., (**2002**). "Probability density functions for hyperbolic and isodiachronic locations," J. Acoust. Soc. Am. **112**, 3046–3052.

Wahlberg, M., Møhl, B., and Madsen, P. T. (**2001**). "Estimating source position accuracy of a large-aperture hydrophone array for bioacoustics," J. Acoust. Soc. Am. **109**, 397–406.

Wald, A. (**1947**). *Sequential Analysis* (Wiley, New York).

Watkins, W. A., Daher, M. A., DiMarzio, N. A., Samuels, A., Wartzok, D., Fristrup, K. M., Howey, P. W., and Maiefski, R. R. (**2002**). "Sperm whale dives tracked by radio tag telemetry," Marine Mammal Sci. **18**, 55–68.

Watwood, S. L., Miller, P. J. O., Johnson, M., Madsen, P. T., and Tyack, P. L. (**2006**). "Deep-diving foraging behaviour of sperm whales (*Physeter macrocephalus*)," J. Anim. Ecol. **75**, 814–825.

White, P. R., Leighton, T. G., Finfer, D. C., Prowles, C., and Baumann, O. (**2006**). "Localisation of sperm whales using bottom-mounted sensors," Appl. Acoust. **67**, 1074–1090.

Whitehead, H. (**2003**). *Sperm Whales: Social Evolution in the Ocean* (Chicago University Press, Chicago).

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

Whitehead, H., and Weilgart, L. (**1991**). "Patterns of visually observable behavior and vocalizations in groups of female sperm whales," Behaviour **118**, 275–296.

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 macreocephalus*)," IEEE J. Ocean. Eng. **28**, 13–28.

Zimmer, W. M. X., Madsen, P. T., Teloni, V., Johnson, M. P., and Tyack, P. L. (**2005a**). "Off-axis effects on the multipulse structure of sperm whale usual clicks with implications for sound production," J. Acoust. Soc. Am. **118**, 3337–3345.

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.