Available online at www.sciencedirect.com
ScienceDirect
Applied Acoustics 67 (2006) 1187–1201
applied acoustics
www.elsevier.com/locate/apacoust

Track of a sperm whale from delays between direct and surface-reflected clicks

Eva-Marie Nosal *, L. Neil Frazer

Department of Geology and Geophysics, School of Ocean and Earth Science and Technology, University of Hawaii at Manoa, 1680 East-West Road Post 813, Honolulu, HI 96822, USA

Received 21 January 2006; received in revised form 18 April 2006; accepted 21 May 2006
Available online 13 July 2006

Abstract

One dataset made available for the 2nd international workshop on detection and localization of marine mammals using passive acoustics featured a single sperm whale recorded for 25 min on five widely spaced, bottom-mounted hydrophones in the Tongue of the Ocean. In this paper, we track the whale using a model-based method that relies on the difference in arrival times along direct and surface-reflected propagation paths. Four receivers are required to estimate positions in three-dimensions. Details of the method are presented, and tracks are estimated using an isospeed and a depth-dependent sound speed profile (SSP). Depth estimates for the isospeed SSP are about 50–100 m shallower than for the depth-dependent SSP, and horizontal positions are similar. Performance estimates indicate that the depth-dependent SSP results are more accurate, with estimated depths of 650–760 m and average vertical and horizontal swim speeds (not accounting for current) of 0.30 m/s and 2.40 m/s, respectively.
© 2006 Elsevier Ltd. All rights reserved.

Keywords: Marine mammals; Passive acoustics; Tracking; Localization

1. Introduction

Recordings of sperm whale vocalizations in the Tongue of the Ocean were made available to participants of the 2nd international workshop on detection and localization of

* Corresponding author. Tel.: +1 808 956 6082; fax: +1 808 956 5154.
E-mail addresses: nosal@hawaii.edu (E.-M. Nosal), neil@soest.hawaii.edu (L. Neil Frazer).

0003-682X/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.apacoust.2006.05.005

marine mammals using passive acoustics. The datasets are from March 23 to 30, 2002 and were prepared by the naval undersea warfare center (NUWC). In the March 23 dataset, a single sperm whale is clearly heard on all five bottom-mounted hydrophones for the full 25 min of recordings.

As they are the predominant vocalization present in the recordings, this work concentrates on the so-called regular (or usual) clicks emitted by sperm whales during deep dives [1]. Regular clicks have inter-click intervals (ICIs) of 0.5–2 s [2,3], duration of about 10–20 ms [2], and energy from below 100 Hz to above 20 kHz [4,5]. Due to these characteristics, as well as the deep-water environment and bottom-mounted hydrophones used for the recordings, direct and surface-reflected arrivals can be easily identified. Specifically, reflected arrivals come shortly after their associated direct arrivals and they have lower amplitude, less power at high frequencies, and slightly longer duration. Moreover, the effect of source-receiver spacing on the time delay between corresponding direct and surface-reflected rays (DRTD) is clearly audible; DRTDs decrease with increasing source/receiver separation. For a moving sperm whale DRTDs vary between receivers as well as with time on a single receiver. Cato [6] and Aubauer et al. [7] explain this effect for an isospeed SSP.

Motivated by this dataset, we developed and implemented a tracking method that relies entirely on DRTDs. DRTDs have previously been exploited for localization [6–10], but they have typically been used with isospeed SSPs to establish range, and not (to our knowledge) to estimate a three-dimensional track using widely spaced receivers, as is done here. A ray-tracing model that accounts for the depth-dependent sound-speed SSP (SSP) gives different, and presumably more accurate, estimates than an isospeed SSP. We also give approximate error maps for depth and x–y coordinates of location. Unfortunately, no visual or tagging data are available to verify the estimated track of the sperm whale. Nevertheless, the estimated track is consistent with other observations of sperm whale behavior, which lends confidence to our predictions.

2. Methods

Before giving the details of the method, we first provide a general overview, noting that at least four receivers are required for the localizations. Signals are sub-divided into short time intervals, and a list of candidate source depths is created. Each time interval and depth is processed separately. At each receiver, the DRTD is established by a click detection scheme. These measured DRTDs are compared to modeled DRTDs to estimate the horizontal separation of source and receiver. This separation defines (the radius of) a circle centered at the receiver. If the search has been conducted at the correct source depth, the arrival times have been accurately determined, and the environment has been perfectly modeled, all receiver circles intersect at a single point, which is the position of the source. In most cases, however, the circles do not intersect at a single point, and a point of best agreement must be determined. This is accomplished by creating a likelihood surface (a probabilistic indicator of source location sometimes referred to as an ambiguity surface) for each receiver; the surfaces are assigned value 1 along the circles (highest probability) and decay according to a Gaussian weighting away from the circle, both inward and outward. Likelihood surfaces for each receiver are averaged to create a total likelihood surface at the current search depth. This process is repeated for all candidate source depths, and the point with maximum likelihood is declared the overall estimated source

position at that time. Only positions at times that give sufficiently large likelihood are retained.

2.1. Preliminaries

The signal at each receiver is sub-divided into short time intervals, typically several tens of seconds long, which can overlap. Two factors are considered when choosing interval lengths. First, longer intervals contain more clicks, which helps to reduce errors in estimated DRTDs. For intervals that are too long, however, movement of the whale may result in significant variation of the DRTDs within the interval. Through trial and error, 20 s intervals (which typically contained between 10 and 25 clicks) were chosen as a good compromise for the workshop dataset. A 15 s overlap was used since it gave good time resolution for the track while keeping run-times reasonably low.

Next, a look-up table of predicted DRTDs as a function of range for all receivers and candidate source depths is created. Hydrophone positions (Table 1) were provided by NUWC. All phones were 17 ft off the bottom except K, which was 18 ft off the bottom. In this work, the Gaussian beam acoustic propagation model BELLHOP [11] was used to model the environment and create this table. The range list varied from 5 m to 10 km in 5 m increments. Since the hydrophones were all within 7.5 km of one another, this allowed for searches several kilometers beyond the boundary of the receiver array. Candidate depths covered the entire water column at 10 m resolution. Two sound speed profiles (SSP) were used. One was an isospeed SSP with a sound-speed of 1510 m/s; the other was a depth-dependent SSP, the average historical SSP from the Tongue of the Ocean for March, taken from the Generalized Digital Environment Model [12] (Fig. 1). Fig. 2 shows modeled DRTDs as a function of horizontal separation for one hydrophone and three candidate source depths.

2.2. Detecting and classifying arrivals to establish DRTDs

As mentioned in the introduction, the characteristics of the source and the environment make it easy to identify direct and surface-reflected arrivals in the time series. Short-duration calls result in no overlap between associated direct and surface-reflected arrivals, and their broadband nature can be used to reduce noise (see next paragraph). The ICI is usually long enough so that a reflected arrival precedes the direct arrival from the next click. The deep-water environment reduces complications from multiple arrivals, and bottom-mounted hydrophones mean that bottom reflections arrive immediately after direct signals, so they are not confused with surface reflections. Refer to Fig. 3 for a waveform of a typical sequence of direct and reflected arrivals. Direct arrivals are high amplitude

Table 1
Hydrophone positions provided by NUWC

Hydrophone	x-Position (m)	y-Position (m)	Depth (m)
G	10658.04	−14953.63	1530.55
H	12788.99	−11897.12	1556.14
I	14318.86	−16189.18	1553.58
J	8672.59	−18064.35	1361.93
K	12007.50	−19238.87	1522.54

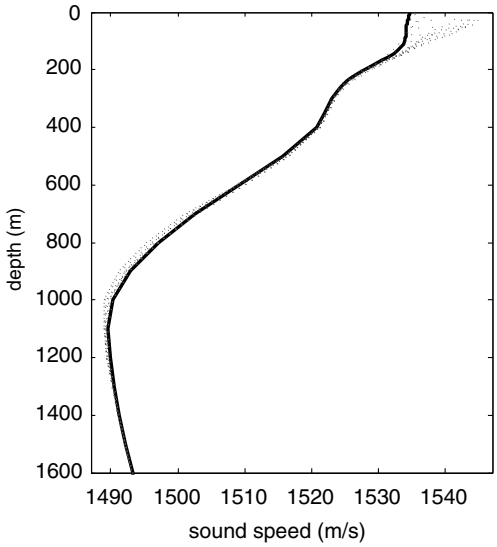


Fig. 1. Historical SSPs for the region. The SSP for March, which is when the data were collected, is shown in bold and was used to model DRTDs.

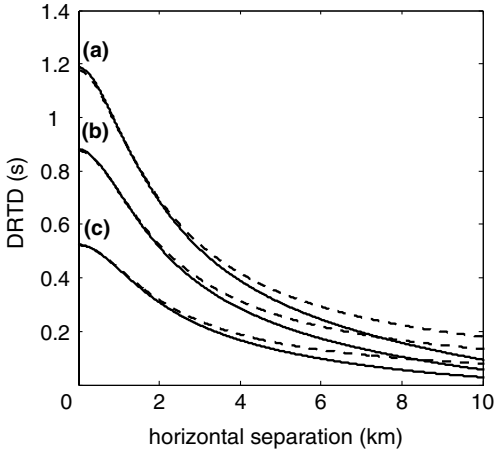


Fig. 2. Modeled DRTDs for hydrophone H as a function of horizontal separation for source depths of: (a) 890 m; (b) 670 m, and (c) 400 m. Solid lines are for the depth-dependent SSP for March; dashed lines are for an isospeed SSP of 1510 m/s.

and quite evenly spaced, with reflected arrivals between them. In some cases the direct and reflected arrivals are not so clear (see Fig. 4(a)).

To improve the detection process for such difficult cases, a spectrogram method was employed. The spectrogram method exploits the broadband nature of the sperm whale clicks. First, a complex spectrogram is created from the hydrophone signal, which has a sampling frequency of 48 kHz, via a short-time Fourier transform. A high-pass filter is

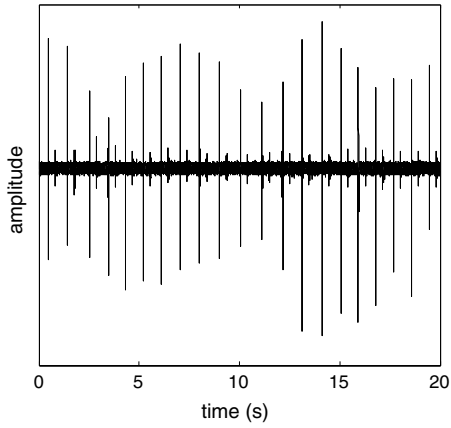


Fig. 3. Waveform at hydrophone J for the data segment beginning at 170 s. Direct arrivals have large relative amplitude. Surface reflections come between the direct arrivals and have smaller relative amplitude.

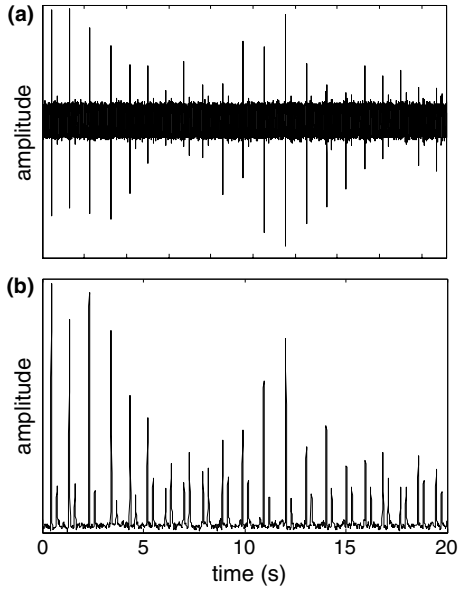


Fig. 4. (a) Waveform at hydrophone K for the data segment beginning at 170 s. Noise completely covers the reflected arrivals. (b) After applying the spectrogram method, noise is significantly reduced and reflected arrivals are clearly identifiable.

applied to the time series of each frequency channel in the spectrogram. Since sperm whale clicks are less than 25 ms in duration, the filter cutoff is set at 40 Hz. Filtering is done in the frequency domain using half of a Hanning window to roll-off with an 80 Hz transition bandwidth. This reduces slowly varying sounds, such as tonal noise from equipment or boats. After filtering, magnitudes are taken of the filtered spectrogram points, and each frequency channel is divided by the mean of the entire channel; this de-emphasizes the

lower frequencies that have more background noise. Finally, the frequency channels are summed to give a time signal with lower background noise and coarser time resolution than the original signal. The time resolution of the channel sum depends on the duration and overlap of the windows used in the discrete FFT transforms. Fig. 5 shows the spectrogram and the processed spectrogram for the signal in Fig. 4(a) using 512-point Hanning windows with 256-point overlap. The channel sum (hereinafter referred to as the filtered signal) is shown in Fig. 4(b).

Peaks in the filtered signal are classified as direct arrivals if they have amplitude greater than some threshold percent of the mean of the three largest amplitudes in the current time interval. For the results presented here, a 50% threshold was used. This was chosen by trial and error via visual inspection of signals and detected arrivals in numerous cases. Too high of a threshold caused many direct arrivals to be missed; too low a threshold caused some reflected arrivals to be incorrectly classified as direct arrivals.

The peak with maximum amplitude arriving between two classified direct arrivals was classified as the surface-reflected arrival corresponding to the direct arrival preceding it. Since direct arrivals were not always entirely impulsive (i.e., they had non-zero duration), and since the bottom-reflected arrival could sometimes be seen immediately following the direct arrival, care was taken not to look for the surface reflection too soon after the direct arrival; arrivals within 100 ms of the direct arrival were not included in the search for surface reflections. Further, since the surface-reflected arrivals were expected to have

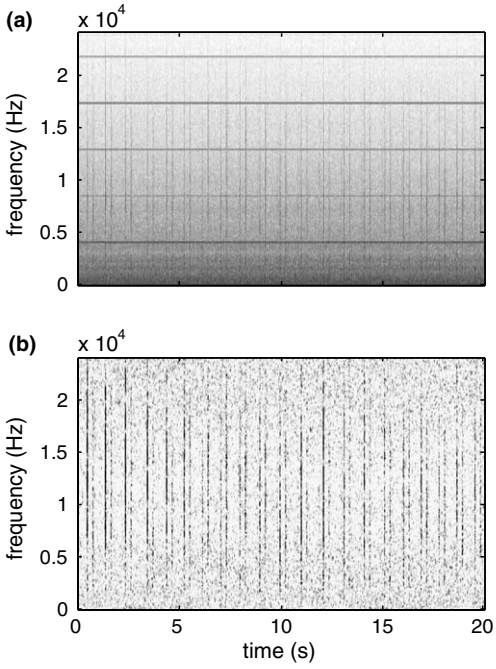


Fig. 5. Spectrogram of the signal from Fig. 4(a) created using 512-point Hanning windows with 256-point overlap. (a) Original spectrogram; and (b) after high-pass filtering and dividing each frequency channel by its mean. Note in (b) that the tonals have been removed, and the low frequencies (with significant background noise) have been de-emphasized.

amplitudes between about 10% and 75% of the amplitude of the corresponding direct arrival, pairs that did not meet these expectations were discarded. Having classified direct and surface-reflected arrivals pairs, DRTDs were computed from their relative arrival times. The median of all resulting times was declared the representative DRTD for the current time window and receiver. A median was taken instead of a mean, since misclassifications can result in large outliers that significantly affect the mean.

In some cases, DRTDs could not be established on all receivers. This occurred at times when the whale was silent, as well as for various cases when the detection scheme failed. This included cases with very poor signal-to-noise ratios and those in which a surface-reflected signal arrived immediately before, at the same time as, or after the subsequent direct arrival (due to short ICIs). Time intervals in which a DRTD could not be established on at least four hydrophones were discarded.

2.3. Creating likelihood surfaces

The following steps are repeated for all candidate source depths. The horizontal separation with modeled DRTD closest to the measured DRTD is found for each receiver. To create a likelihood surface, a grid is created that covers the horizontal plane of interest. For the workshop dataset, the grid range used was 7000 m to 16000 m N–S and −21000 m to −10000 m E–W, with 10 m resolution in both directions. For each receiver, the likelihood value is a function of each grid point's radial distance from the perimeter of a circle centered at the receiver with radius given by the horizontal separation corresponding to that receiver. A Gaussian weighting function, with standard deviation 500 m, was found to work well for the current dataset. Likelihood surfaces for all receivers on which DRTD could be established are averaged to give the total likelihood surface (with value between 0 and 1) at the current candidate source depth. Fig. 6 shows the likelihood surface at three different depths. When the candidate source depth is too shallow or too deep, the receiver circles do not intersect closely, resulting in lower maxima. The positions and values of the maxima are stored.

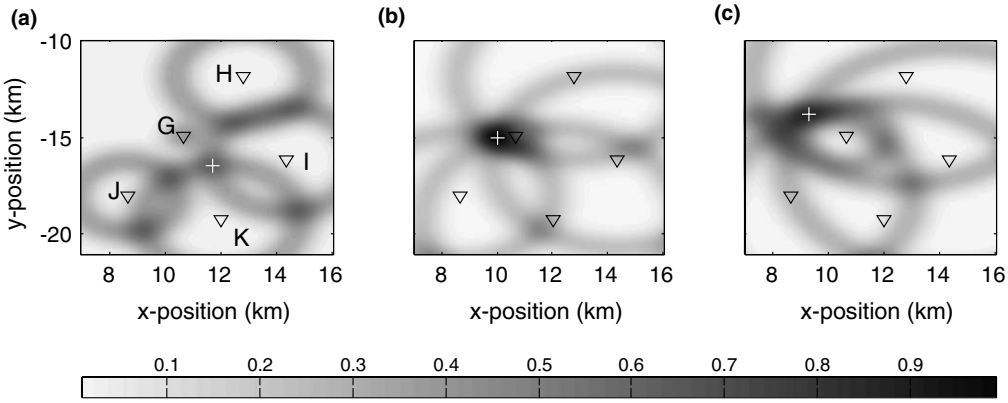


Fig. 6. Likelihood surfaces for the first 20 s of data at depths of (a) 400 m (b) 670 m and (c) 890 m. Triangles indicate receiver locations and are labeled in (a). White crosses mark position estimates, with surface values (a) 0.401 (b) 0.997 and (c) 0.729. The estimated source position, chosen from (b), is (10010 m, −15020 m, 670 m).

After this process has been completed for all candidate source depths, the point with the maximum likelihood value is chosen as the estimated source position. Smaller maximum likelihood values indicate greater uncertainty in the source position. Times with likelihood below a preset level are discarded as having too much error. For the workshop data, a threshold of 0.850 was used.

3. Results

The process was automated by a collection of MATLAB codes. No effort was made to optimize the code for efficiency. Modeling of the environment (calculation of predicted DRTDs) took less than a minute and was done once for each SSP. After this, run times were about three times real-time on a 2.8 MHz Pentium IV; 5 min of data took about 17 min to process. Using coarser time and/or space resolution can reduce run times. Also, a more intelligent search can increase the efficiency of the algorithm (e.g., the swim speed of the whale is limited so it is not necessary to search the entire water column if the position of the whale is established for previous time intervals).

Results are shown in Fig. 7 as estimated position (x-, y-, and depth) versus time. In these results, 13% of all time intervals were eliminated because DRTDs could not be established on four or more receivers. Of the remaining time intervals, 24% were eliminated in the isospeed SSP case and 20% were eliminated in the depth-dependent SSP case because maximum likelihood values were less than 0.850. The mean of the maximum likelihood values for the remaining points is 0.920 for the isospeed SSP and 0.998 for the depth-

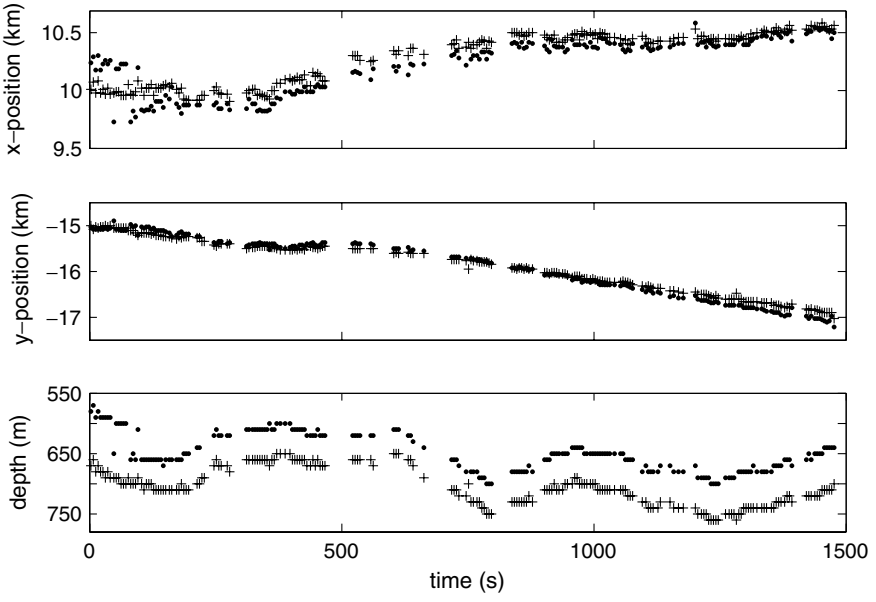


Fig. 7. Estimated track (not smoothed) of the sperm whale with the isospeed SSP (dots) and the depth-dependent SSP (crosses). Times with performance values less than 0.850, and those for which DRTDs could be established on only three (or fewer) receivers, have been eliminated.

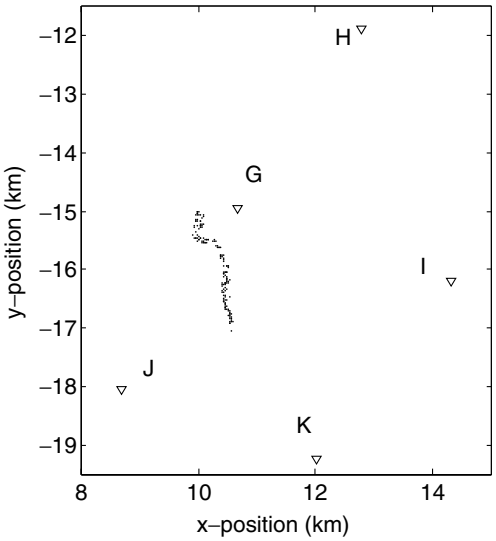


Fig. 8. Estimated two-dimensional track (not smoothed) of the sperm whale with the depth-dependent SSP. Triangles indicate receiver positions.

dependent SSP, suggesting that the depth-dependent SSP results are more accurate. The x- and y- tracks for the isospeed SSP and the depth-dependent SSP are quite similar. The depth tracks are also similar, with the depth-dependent SSP track between 650 m and 760 m and about 100 m deeper than the isospeed track. This difference reiterates [7,8,13] the importance of incorporating the effects of a depth-dependent SSP into methods for

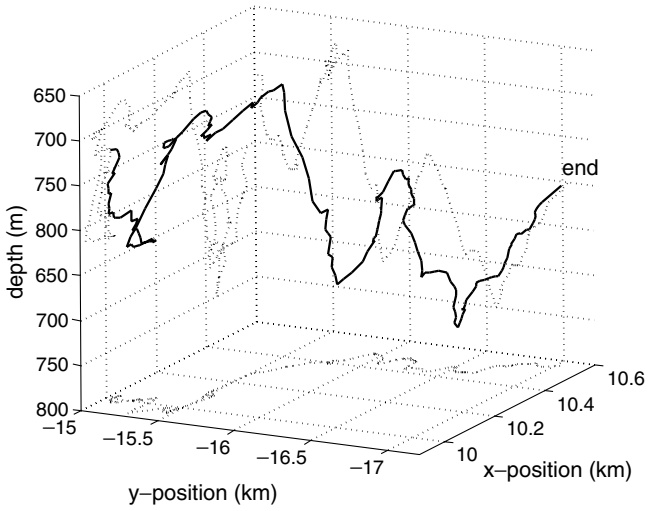


Fig. 9. Smoothed three-dimensional track (solid line) estimated using the depth-dependent SSP. Projections onto the three planes are shown with dashed lines.

tracking marine mammals over long ranges using passive acoustics. In Fig. 8, the track from the depth-dependent SSP is plotted in the x–y plane with the hydrophone positions. Fig. 9 shows the smoothed track from the depth-dependent SSP in three-dimensions. A five-point moving average filter was used along each direction to accomplish the smoothing.

The average swim speed was estimated from the smoothed track by calculating the velocity for each time step and taking the mean over all times. The vertical swim speed was 0.30 m/s and the horizontal swim speed was 2.40 m/s. It is not known what the current was in the area at the time, so the horizontal swim speed relative to the current cannot be estimated.

4. Error estimates

To estimate the error in source position associated with the method presented here, error maps of the array for x- y- and z-directions were created under the assumptions that sources of error are independent, error in one direction is independent of error in the other directions, and errors are normally distributed. As in time-difference of arrival methods [14], there are three main input variables (hence sources of error) associated with the DRTD method: sound speed SSP, measured DRTD, and receiver position. Since no measurement of receiver position error was available, we deal only with errors due to SSP and DRTD.

First, consider errors in the x- and y-directions. For receiver i and position $\bar{p} = (x_p, y_p, z_p)$, depth is fixed to find $\sigma_{r,i}$, the standard deviation in range for receiver i (see Fig. 2). This is a combination of the error due to the SSP, $\sigma_{\mathrm{ssp},r,i}$, and the error due to DRTD, $\sigma_{\mathrm{drtd},r,i}$. Letting t_d be the modeled DRTD between receiver i and $\bar{p}$ for a reference SSP, we approximate $\sigma_{\mathrm{ssp},r,i}$ as the difference between the minimum and maximum ranges corresponding to t_d over all possible SSPs, and $\sigma_{\mathrm{drtd},r,i}$ as one fourth of the difference between the ranges corresponding to $t_d \pm 2\sigma_{\mathrm{drtd}}$ for the reference SSP, where σ_{drtd} is the estimated standard deviation of DRTD measurements. To the circle passing through $\bar{p}$, centered on receiver i, we attach a Gaussian PDF with standard deviation

$$\sigma_{r,i} = \sqrt{\sigma^2_{\mathrm{ssp},r,i} + \sigma^2_{\mathrm{drtd},r,i}}.$$

To estimate error in the x-direction from the combined PDFs at all receivers, each PDF is approximated as locally linear. In other words, the PDF for each receiver is approximated by a ridge whose axis is the line tangent to the corresponding circle at $\bar{p}$. Let θ_i denote the angle from the x-axis to the ith hydrophone, then the combined PDF at point (x, y_p, z_p) is:

$$p(x) \propto \prod_i \exp\left[-\frac{(x-x_p)^2}{2}\frac{\cos^2(\theta_i)}{\sigma^2_{r,i}}\right] = \exp\left[-\frac{(x-x_p)^2}{2}\sum_i \frac{\cos^2(\theta_i)}{\sigma^2_{r,i}}\right]. \quad (1)$$

Normalization is automatic as:

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma_x}\exp\left[-\frac{(x-x_p)^2}{2\sigma_x^2}\right]; \text{ in which } \sigma_x = \left(\sum_i \frac{\cos^2(\theta_i)}{\sigma^2_{r,i}}\right)^{-1/2}. \quad (2)$$

Similarly for error in the y-direction, the combined PDF at point (x_p, y, z_p) is:

$$p(y) = \frac{1}{\sqrt{2\pi}\sigma_y} \exp\left[-\frac{(y-y_p)^2}{2\sigma_y^2}\right]; \text{ in which } \sigma_y = \left(\sum_i \frac{\sin^2(\theta_i)}{\sigma_{r,i}^2}\right)^{-1/2}. \quad (3)$$

Error in the z-direction is handled in a similar manner. For receiver i and position $\bar{p} = (x_p, y_p, z_p)$, range is fixed to find the standard deviations in depth for receiver i: $\sigma_{\text{ssp},d,i}$, $\sigma_{\text{drtd},d,i}$, and $\sigma_{d,i} = \sqrt{\sigma_{\text{ssp},d,i}^2 + \sigma_{\text{drtd},d,i}^2}$ (see Fig. 10). The combined PDF at point (x_p, y_p, z) is simply:

$$p(z) = \frac{1}{\sqrt{2\pi}\sigma_z} \exp\left[-\frac{(z-z_p)^2}{2\sigma_z^2}\right]; \text{in which } \sigma_z = \left(\sum_i \sigma_{d,i}^{-2}\right)^{-1/2}. \quad (4)$$

Based on the width of the processed clicks (about 10 ms), the standard deviation of the DRTD measurements, σ_{drtd}, was set to 5 ms. Fig. 11(a)–(c) shows location error maps (at 700 m depth) obtained using the SSPs from all 12 months (see Fig. 1) as the collection of possible SSPs. Error maps for different depths are similar. For the horizontal directions, error from DRTD measurement only is about double the error from SSP only. In the vertical direction, DRTD measurement errors are similar to SSP errors. Fig. 11(d)–(e) shows error maps obtained when the isospeed SSP is added to the collection of possible SSPs. In this case, errors associated with SSP only are an order of magnitude greater than errors associated with DRTD only. In Fig. 12, the errors associated with the actual tracks are plotted, with the errors calculated for (a) the depth-dependent case using the monthly SSPs and (b) the isospeed case using the isospeed SSP in addition to the monthly SSPs. In all instances, error in the vertical direction is less than error in the horizontal directions; this is expected because DRTD changes more rapidly with depth (Fig. 10) than with range (Fig. 2).

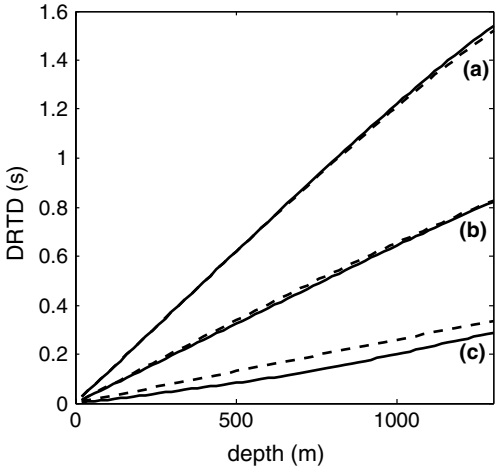


Fig. 10. Modeled DRTDs for hydrophone H as a function of depth for source ranges of: (a) 500 m; (b) 2500 m, and (c) 7500 m. Solid lines are for the depth-dependent SSP for March; dashed lines are for an isospeed SSP of 1510 m/s.

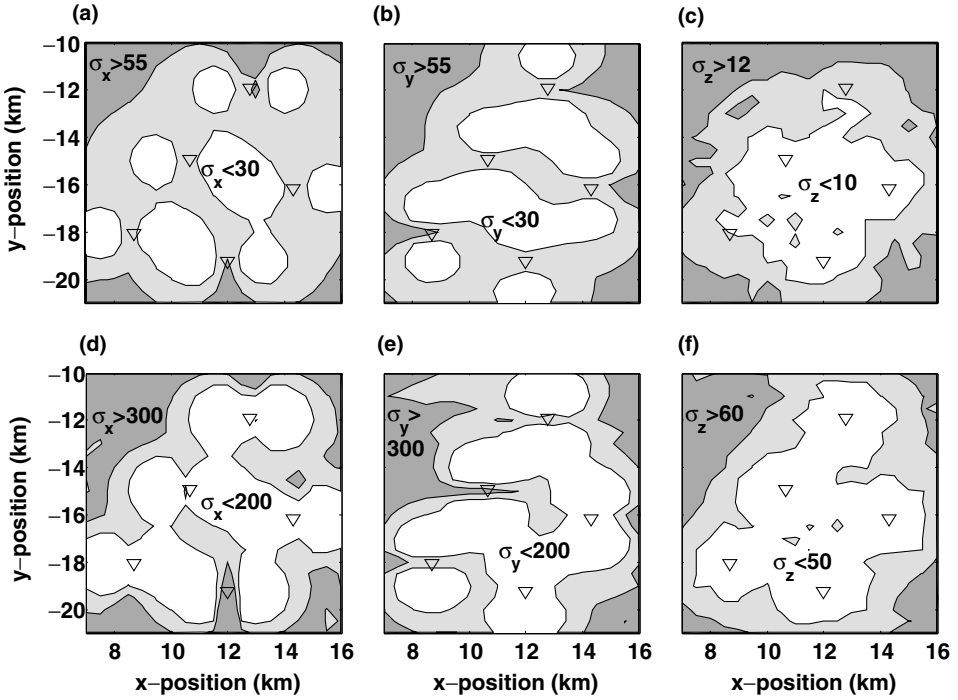


Fig. 11. Contour intervals (values as indicated) of one standard deviation in x-, y-, and z-source position at 700 m depth due to uncertainties in SSPs and DRTDs. (a)–(c) Depth-dependent SSP; (d)–(e) isospeed SSP. Triangles indicate receiver positions.

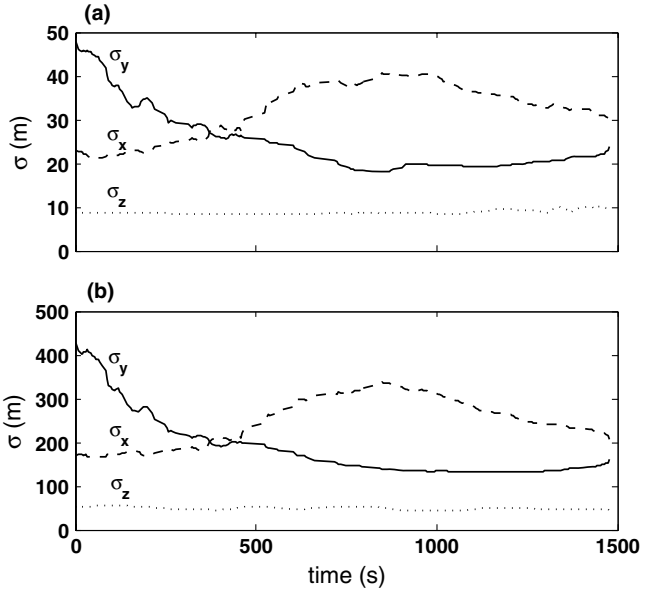


Fig. 12. One standard deviation in x-, y-, and z-source position (as a function of time) along estimated tracks due to uncertainties in SSPs and measurement of DRTDs. (a) Depth-dependent SSP; (b) isospeed SSP.

5. Discussion

Although it cannot be confirmed by tags or sightings, the estimated track is consistent with what is expected for a sperm whale. In particular, sperm whale dives are typically many hundreds of meters deep [4,15–17] with reports of dives in excess of 1000 m [10,18]. Dives may last up to 90 min [2], but are more commonly between about 25 and 50 min [4,16,20,21]. Also, the estimated swim speeds agree with those observed in previous studies [4,16,18,19]. Our error estimates suggest that the track using the depth-dependent SSP is correct to about 100 m in horizontal position and 20 m in depth.

An important advantage of the DRTD method over arrival time difference methods commonly used for marine mammal localization is that it is much less sensitive to synchronization errors in timing between receivers. This is because DRTD measurements are estimated for individual hydrophones, rather than between pairs of hydrophones. Although a comprehensive study of synchronization error was not performed, it is worth noting that a 2.34 s offset between two of the five hydrophones that was (unknown and) present in the original version of the distributed dataset did not significantly affect our predicted track.

Several problems are associated with the DRTD method. First, for near-surface sources, direct and surface-reflected clicks are difficult to distinguish. Shadow zones present another problem for near-surface sources, although this effect would likely occur on only one receiver, and can be overcome for sufficiently large (>5 receiver) arrays. Furthermore, surface roughness associated with gravity waves may have a significant effect on reflected arrival times [9,22], and hence on estimated DRTDs. Uncertainty in receiver location is also a problem, as it is with all localization techniques. Methods to locate the receivers more accurately [23], or to include variable receiver position in the modeling are useful for this [24]. Our detection scheme is for a single animal, but improved schemes that can distinguish calls of individuals [25] might extend its applicability to multiple whales. Finally, it would be prudent to compare and combine the DRTD method with other localization techniques to give more accurate track estimates.

The reader is advised that some of the methods presented here were done with a somewhat "quick and dirty" mentality. This approach was taken because we wanted to test the feasibility of using DRTDs for three-dimensional localization without getting tangled in detail, and there is certainly much room for improvement. For example, the method would benefit from a more sophisticated (and objective) detection and classification scheme [5]. Also, likelihood surfaces should incorporate errors in measurement and modeling instead of using the (empirical and somewhat arbitrary) standard deviation of 500 m. Rather than searching over candidate depths, likelihood volumes could be created in three-dimensions. Further, the receiver log likelihoods (rather than the likelihoods themselves) should be averaged to create overall likelihood surfaces. Among other possible improvements, these things would reduce errors and allow for more accurate error estimates.

6. Conclusions

Recordings of a single sperm whale on five bottom-mounted hydrophones in a deep-water environment were used to track the animal in three-dimensions for 25 min. A model-based method based on the arrival time difference between direct and surface-reflected clicks was used in the tracking and described in detail. All five hydrophones were

used and at least four hydrophones are needed to apply the method. A depth-dependent SSP led to better performance estimates than an isospeed SSP. Run times were about three times longer than real-time, but can be reduced to real-time by decreasing resolution or by using a faster machine. Although we did not have data to verify the track visually or otherwise, it is consistent with sperm whale behavior. Estimated horizontal positions were similar for both SSPs, but depth for the isospeed SSP was about 50–100 m shallower than for the depth-dependent SSP. The estimated depth of the whale varied between 650 m and 760 m for the depth-dependent SSP. The average vertical and horizontal swim speeds were 0.30 m/s and 2.40 m/s, respectively.

Acknowledgements

We gratefully acknowledge NUWC for providing the dataset. Thanks to all workshop organizers, particularly to O. Adam. Thanks also to F. Desharnais and D. Gillepsie, for compiling the dataset results and for general discussion, and to two anonymous reviewers for highly detailed and constructive comments. This work was supported by the Office of Naval Research.

References

[1] Whitehead H, Weilgart L. Patterns of visually observable behavior and vocalizations in groups of female sperm whales. Behavior 1991;118:275–96.
[2] Goold JC, Jones SE. Time and frequency domain characteristics of sperm whale clicks. J Acoust Soc Am 1995;98(3):1279–91.
[3] Jaquet N, Dawson S, Douglas L. Vocal behavior of male sperm whales: why do they click? J Acoust Soc Am 2001;109:2254–9.
[4] Watkins WA, Daher MA, Fristrup KM, Howald TJ, DiSciara GN. Sperm whales tagged with transponders and tracked underwater by sonar. Mar Mammal Sci 1993;9(1):55–67.
[5] Zimmer WMX, Tyack PL, Johnson MP, Madsen PT. Three-dimensional beam pattern of regular sperm whale clicks confirms bent-horn hypothesis. J Acoust Soc Am 2005;117:1473–85.
[6] Cato D. Simple methods of estimating source levels and locations of marine animal sounds. J Acoust Soc Am 1998;104(3):1667–78.
[7] Aubauer R, Lammers M, Au W. One-hydrophone method of estimating distance and depth of phonating dolphins in shallow water. J Acoust Soc Am 2000;107(5):2744–9.
[8] Thode A. Three-dimensional passive acoustic tracking of sperm whales (*Physeter macrocephalus*) in ray-refracting environments. J Acoust Soc Am 2005;18(6):3575–84.
[9] Skarsoulis EK, Kalogerakis MA. Ray-theoretic localization of an impulsive source in a stratified ocean using two hydrophones. J Acoust Soc Am 2005;118(5):2934–43.
[10] Zimmer WMX, Johnson MP, D'Amico A, Tyack P. Combining data from a multisensor tag and passive sonar to determine the diving behavior of a sperm whale (*Physeter macrocephalus*). IEEE J Oceanic Eng 2003;28(1):13–28.
[11] Porter MB, Liu YC. Finite-element ray tracing. In: Lee D, Schultz MH, editors. Proc Int Conf Theor Comp Acoust. Singapore: World Scientific; 1994. p. 947–56.
[12] Current website. Available from: https://128.160.23.42/gdemv/gdemv.html.
[13] Chapman DMF. You can't get there from here: shallow water sound propagation and whale localization. Can Acoust 2004;32(2):167–71.
[14] Wahlberg M. Comparing a linear with a non-linear method for acoustic localization. Can Acoust 2004;32(2):125–31.
[15] Clarke MR, Martins HR, Pascoe P. The diet of sperm whales, *Physeter macrocephalus Linnaeus* 1758, off the Azores. Philos Trans R Soc London B 1993;339:67–82.
[16] Papastavrou V, Smith SC, Whitehead H. Diving behavior of the sperm whale, *Physeter macrocephalus*, off the Galapagos Islands. Can J Zool 1989;67:839–46.

[17] Lockyer C. Observations on diving behaviour of the sperm whale, *Physeter catodon*. In: Angel M, editor. A voyage of discovery. Oxford: Pergamon; 1977. p. 591–609.
[18] Wahlberg M. The acoustic behaviour of diving sperm whales observed with a hydrophone array. J Exp Mar Biol Ecol 2002;281:53–62.
[19] Watkins WA, Daher MA, DiMarzio NA, Samuels A, Wartzok D, Fristrup KM, et al. Sperm whale dives tracked by radio tag telemetry. Mar Mammal Sci 2002;18(1):55–68.
[20] Gordon JCD, Steiner L. Ventilation and dive patterns in sperm whales, *Physeter macrocephalus*, in the Azores. Rep Int Whal Comm 1992;42:561–5.
[21] Jaquet N, Dawson S, Slooten E. Seasonal distribution and diving behavior of male sperm whales off Kaikouta: Foraging implications. Can J Zool 2000;78:407–19.
[22] Godin OA, Fuks IM. Travel-time statistics for signals scattered at a rough surface. J Acoust Soc Am 1989;85:83–9.
[23] Wahlberg M, Madsen PT, Møhl B. Estimating source position accuracy of a large aperture hydrophone array for bioacoustics. J Acoust Soc Am 2001;109(1):397–406.
[24] Michalopoulou Z-H, Ma X. Source localization in the Haro Strait primer experiment using arrival time estimation and linearization. J Acoust Soc Am 2005;118(5):2924–33.
[25] Thode A, Mellinger DK, Stienessen S, Martinez A, Mullin K. Depth-dependent acoustic features of diving sperm whales (*Physeter macrocephalus*) in the Gulf of Mexico. J Acoust Soc Am 2002;112(1):308–21.