MIT Open Access Articles

Using a coherent hydrophone array for observing sperm whale range, classification, and shallow-water dive profiles

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.


As Published: http://dx.doi.org/10.1121/1.4874601

Publisher: Acoustical Society of America (ASA)

Persistent URL: http://hdl.handle.net/1721.1/97630

Version: Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

Terms of Use: Article is made available in accordance with the publisher’s policy and may be subject to US copyright law. Please refer to the publisher’s site for terms of use.
Using a coherent hydrophone array for observing sperm whale range, classification, and shallow-water dive profiles

Duong D. Tran, Wei Huang, Alexander C. Bohn, and Delin Wang

Department of Electrical and Computer Engineering, Northeastern University, 360 Huntington Avenue, Boston, Massachusetts 02115

Zheng Gong and Nicholas C. Makris

Department of Mechanical Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139

Purnima Ratilala[a]

Department of Electrical and Computer Engineering, Northeastern University, 360 Huntington Avenue, Boston, Massachusetts 02115

(Received 17 September 2013; revised 20 February 2014; accepted 11 April 2014)

Sperm whales in the New England continental shelf and slope were passively localized, in both range and bearing, and classified using a single low-frequency (<2500 Hz), densely sampled, towed horizontal coherent hydrophone array system. Whale bearings were estimated using time-domain beamforming that provided high coherent array gain in sperm whale click signal-to-noise ratio. Whale ranges from the receiver array center were estimated using the moving array triangulation technique from a sequence of whale bearing measurements. Multiple concurrently vocalizing sperm whales, in the far-field of the horizontal receiver array, were distinguished and classified based on their horizontal spatial locations and the inter-pulse intervals of their vocalized click signals. The dive profile was estimated for a sperm whale in the shallow waters of the Gulf of Maine with 160 m water-column depth located close to the array’s near-field where depth estimation was feasible by employing time difference of arrival of the direct and multiply reflected click signals received on the horizontal array. By accounting for transmission loss modeled using an ocean waveguide-acoustic propagation model, the sperm whale detection range was found to exceed 60 km in low to moderate sea state conditions after coherent array processing.

© 2014 Acoustical Society of America. [http://dx.doi.org/10.1121/1.4874601]

PACS number(s): 43.30.Pc, 43.30.Sf [AMT] Pages: 3352–3363

I. INTRODUCTION

Sperm whales (Physeter macrocephalus) in the Northwest Atlantic during spring and summer are concentrated along continental slope regions from the Mid-Atlantic Bight to south of Georges Bank (Whitehead et al., 1992) and the Scotian shelf edge. While foraging, sperm whales perform deep dives lasting from several minutes to more than an hour (Watwood et al., 2006), emitting short-duration broadband clicks with frequencies ranging from several hundred hertz to more than 30 kHz (Madsen et al., 2002b; Weilgart and Whitehead, 1988). Each click exhibits a multi-pulse structure (Møhl, 2001; Norris and Harvey, 1972; Zimmer et al., 2004) arising from reflections of the acoustic signal generated by the phonic lips off the frontal and distal air sacs bounding the spermaceti organ of a sperm whale. The inter-pulse interval (IPI) provides a measure of the length of the spermaceti organ that has been shown to be strongly correlated with the size of a sperm whale individual (Antunes et al., 2010; Gordon, 1990; Growcott et al., 2011; Mathias et al., 2009; Miller et al., 2013; Rhinelander and Dawson, 2004; Teloni et al., 2007). Here we show that it is possible to distinguish and classify multiple vocalizing sperm whale individuals located in the far-field of a single, densely sampled, towed horizontal coherent hydrophone array system using the instantaneous sperm whale position estimates in both range and bearing, and the IPIs of the vocalized click signals. Most studies of the vocalization behavior and dive profile of sperm whales have been confined to deep continental slope environments bounding the Pacific and Atlantic ocean (Jacquet et al., 2001; Madsen et al., 2002b; Mathias et al., 2012; Watwood et al., 2006). Here we provide estimates of the three-dimensional (3D) dive profile of a sperm whale individual the vocalizations of which were opportunistically recorded in the shallow water environment of the Gulf of Maine with roughly 160 m water-column depth during a sea test of a newly developed, densely sampled, towed horizontal coherent receiver array system in May 2013.

Localization of an acoustic source, such as a vocalizing sperm whale, in the far-field of a single, densely sampled, towed horizontal coherent hydrophone array system is often a two-stage process. First, the bearing or horizontal direction of arrival of the acoustic signal is determined by time-delay analysis or beamforming of the signals received on the individual hydrophone elements of the array. Second, the range

[a]Author to whom correspondence should be addressed. Electronic mail: purnima@ece.neu.edu

PACS number(s): 43.30.Pc, 43.30.Sf [AMT] Pages: 3352–3363
or horizontal distance of the acoustic source from the receiver array center is determined by tracking changes in the bearing of a series of acoustic emissions over time (Gong et al., 2013; Nardone et al., 1984; Oshman, 1999; Ristic et al., 2004; Yardim et al., 2011). Short-aperture, towed horizontal coherent hydrophone array systems have been previously used to record vocalizations from sperm whales. However, the coherent receiver array data have only been applied to determine the bearing of a sperm whale. No subsequent range estimates have been made based solely on the coherent receiver array measurements. Consequently, coherent array gain of densely sampled hydrophone array systems has not been previously exploited for range localization of sperm whales. In Teloni (2005), a 128-element horizontal coherent hydrophone array system of the NATO Undersea Research Center (NURC) with an array aperture length of 11.6 m was employed to determine sperm whale vocalization bearings and to separate whale clicks from different azimuthal directions, but no range estimates or range analysis were provided. In Zimmer et al. (2004), click data from a single sperm whale acquired using the same NURC receiver array were used to determine the bearings of the whale individual. The sperm whale was primarily tracked using a digital tag (DTAG) attached to its body consisting of a hydrophone used to record sounds directly from the whale (Zimmer et al., 2004). The sperm whale range to the receiver array center was determined from click travel time difference between the DTAG hydrophone and the towed array hydrophones (Zimmer et al., 2004).

Here we localize multiple sperm whales in the far-field of a single low-frequency (<2.5 kHz), densely sampled, towed horizontal coherent hydrophone array system, providing estimates of both range and bearing for each sperm whale. Because no other acoustic sensors were available to us apart from the towed horizontal receiver array system, the whale ranges were estimated from their bearing-time trajectories. A review of methods that can be applied to passively estimate the range of an acoustic source from a single, densely sampled, towed horizontal coherent receiver array is provided in Sec. I of Gong et al. (2013). Here we employ the moving array triangulation technique developed in Gong et al. (2013) to estimate sperm whale ranges from the measured click bearings. This technique combines bearing measurements from spatially separated apertures of the towed horizontal coherent receiver array and employs the conventional triangulation ranging algorithm for localizing a source that may be in the near- or far-field of the array. Because data from only a single towed horizontal coherent receiver array are used here to remotely and passively localize both the range and bearing of sperm whales and to classify them, the methods and results developed here are highly relevant and can be directly applied to address the feasibility of monitoring sperm whales with other towed horizontal coherent receiver array systems, such as those employed in naval and geophysical applications, where it may be important and necessary to remotely sense marine mammal activity from long ranges. An advantage of bearings-only range localization with a densely sampled, towed horizontal coherent receiver array system is that no additional information about the environment, such as bathymetry or sound speed profile, is needed to estimate source range in the far-field of the array.

Other approaches for localizing sperm whales include (a) hyperbolic ranging with a small network of single hydrophones (Baggenstoss, 2011; Tiemann and Porter, 2003; Watkins and Schevill, 1972) and (b) time-delay measurement of click reflections from ocean boundaries acquired with a single hydrophone or a sparse array of hydrophones (Mathias et al., 2013; Mathias et al., 2012; Nosal and Frazer, 2006; Skarsoulis and Kalogerakis, 2006; Thode, 2004; Tiemann et al., 2006; Wahlberg, 2002). Here we utilize time difference of arrivals of the sperm whale direct and multiple bottom and surface reflected click [multiple-reflection based time difference of arrival (MR-TDA)] signals after beamforming to estimate the depth and hence the dive profile of a sperm whale in shallow waters of the Gulf of Maine with 160 m water-column depth. This sperm whale individual’s horizontal range \( r = 1 \text{ km} \) was very close to the array’s near-field distance \( r_N \) \( (r_N \leq L^2/\lambda \approx 750 \text{ m}) \), where \( L \) is the array aperture length and \( \lambda \) is the wavelength) making it possible to estimate its depth. Depth estimation for acoustic sources at long ranges, in the far-field \( (r \gg L^2/\lambda) \) of a single, horizontal coherent receiver array system is challenging because the acoustic wavefield received by the array is multi-modal in nature having undergone many surface and bottom bounces in a random ocean waveguide making the received field less sensitive to the source’s depth location, except in the endfire direction of the horizontal array.

II. METHOD: EXPERIMENTAL DATA COLLECTION AND ANALYSIS

A newly developed, densely sampled, towed horizontal linear hydrophone array system funded by the National Science Foundation and the Office of Naval Research was deployed and tested in the continental slope region south of Cape Cod between 500 and 2000 m water depth on May 13 (site B in Fig. 1) and in the Gulf of Maine in 150–180 m water depth on May 14 and 15 (site A in Fig. 1). Passive acoustic data were collected on a sub-aperture of the array with \( N = 32 \) elements having an inter-element spacing of 0.75 m. The hydrophone elements, each having \(-188 \text{ dB re } \mu \text{Pa/V}\) sensitivity were sampled at 5 kHz with 24-bit digital resolution. The array was towed by the research vessel Endeavor at various speeds between 1 and 4 kn. The water-column sound speed at the experiment sites were monitored using a conductivity-temperature-depth (CTD) sensor and expendable bathythermographs (XBT). The measured sound speed profiles are shown in Fig. 2 for the two test sites. The array depth was maintained between 60 and 80 m near the thermocline in the Gulf of Maine, and varied between 10 and 50 m at the continental slope region.

To investigate the presence of sperm whale clicks, time-frequency spectrograms of the received signal on each hydrophone were first calculated, and the spectrograms incoherently averaged across several hydrophones were obtained. Sperm whale clicks were consistently present in all 75 min of passive acoustic recordings on May 13 at the continental
The maximum frequency of the acoustic data recorded by the receiver array system here is 2.5 kHz. Our analysis of sperm whale clicks is therefore limited to the low-frequency component of the click signals in the several hundred hertz range that is more omnidirectional (Mathias et al., 2013; Tiemann et al., 2006; Zimmer et al., 2004) and suffer less transmission loss. In contrast, the sampling frequencies of acoustic systems in previous sperm whale studies were significantly higher, by at least a factor of three (Mathias et al., 2013; Tiemann et al., 2006) to over ten times that used in this study to provide a more complete coverage of the bandwidth of the sperm whale click that can extend to 30 kHz (Madsen et al., 2002a; Mathias et al., 2012; Teloni, 2005; Thode, 2004; Watwood et al., 2006; Zimmer et al., 2004).

A. Sperm whale localization in horizontal range and azimuthal bearing

1. Determining click arrival time and azimuthal bearing

The relative horizontal azimuthal direction or relative bearing $\beta$ of each sperm whale click, measured from array broadside, was next estimated using time-domain delay-and-sum beamforming (Johnson and Dudgeon, 1993). The pressure-time series data from each hydrophone of the array within roughly 13 s duration was next converted to 2D beam-time data by steering the array in 400 azimuthal directions equally spaced from $-1$ to 1 in $\sin \beta'$, where $\beta'$ is the azimuthal angle measured from array broadside. An angle of $\sin \beta' = -1$ corresponds to the back endfire direction and $\sin \beta' = 1$ corresponds to the forward endfire direction. The relative azimuthal direction and time of arrival of each sperm whale click were determined from the local peak energy levels of the 2D beam-time data. In general, the sperm whale clicks after high-pass filtering and beamforming stood between 10 and 35 dB above the background. A detection threshold of 10 dB above the background was applied in the local peak detection to reduce the false alarm rate.

Spectrogram and time-series examples of the beamformed received click trains are shown in Figs. 3 and 4. Because each sperm whale click contains multiple sharp pulses highly localized in time with a width of less than 1 ms per pulse (see Fig. 5), the click signal resembles the output of a matched filter. This enables high resolution beamforming in the time domain because coherent addition of the pulses across all hydrophones decorrelates within a small time lag of roughly 1/8 ms, corresponding to bearing estimation accuracies of approximately 1.7$^\circ$ at array broadside and 8$^\circ$ near array endfire. In contrast, the array angular resolution is much broader, roughly $\lambda/(L \cos \beta') = 3.7^\circ$ at broadside and $2\sqrt{0.886\lambda/L} = 22.7^\circ$ at endfire from planewave beamforming of a time-harmonic signal, where $\lambda$, $L$, and $\beta'$ are, respectively, the wavelength, array aperture and azimuthal direction from array broadside (Johnson and Dudgeon, 1993; Makris et al., 1995) for the given array aperture $L = 23.25$ m at a frequency of 2 kHz. Examples of the beam pattern obtained from beamforming two distinct sperm whale clicks, one located near array broadside and the other near array endfire are shown in Fig. 6. The direction of arrival is clearly distinguishable since the main lobe stands more than 8 dB above the grating lobe in both cases, mitigating any potential effect of spatial aliasing.

The estimated relative bearings $\beta'$ measured with respect to array broadside were then converted to absolute bearings $\beta$, measured from the array center with respect to true North by correcting for the corresponding array heading.
measurement \( z \). To resolve the left-right ambiguity inherent in linear receiver array measurement of a source bearing, the left and right absolute bearing sequences were statistically correlated to the measured array headings. The true bearing sequence was selected to be the one with the smaller correlation coefficient because the ambiguous bearing sequence closely follows the array heading changes as shown in Fig. 2 of Gong et al. (2013). When the array is steered in the azimuth of sperm whale clicks, the array gain (Kay, 1998; Urick, 1983) obtained from coherent addition of the click signals measured across all \( N = 32 \) hydrophones can enhance the signal-to-noise ratio by approximately \( 10 \log_{10} N \approx 15 \) dB over that of a single hydrophone [compare beamformed click signals in Fig. 4(C) with single hydrophone measured click signals].

---

FIG. 3. (Color online) (A) Spectrogram of a series of sperm whale echolocation clicks recorded at frequencies up to 2.5 kHz in the Gulf of Maine on May 14 starting at 17:19:07 EDT. The spectrogram was calculated using a short-time Fourier transform with window size 256 and 75% overlap. (B) Beamformed pressure time series of the clicks, bandpass filtered between 1500 and 2500 Hz. (C) Beamformed pressure time series plotted in decibel (dB) scale. The solid curve with error bars shows the mean and standard deviation of beamformed background ambient noise level in the 1500–2500 Hz band, estimated from regions outside of clicks.

FIG. 4. (Color online) (A) Spectrogram of three consecutive slow clicks followed by a train of echolocation clicks recorded at frequencies up to 2.5 kHz in the Gulf of Maine on May 14 starting at 17:16:15 EDT. (B) Beamformed pressure time series of the clicks, bandpass filtered between 500 and 2500 Hz. (C) Beamformed pressure time series plotted in dB scale. The solid curve with error bars shows the mean and standard deviation of beamformed background ambient noise level in the 500–2500 Hz band, estimated from regions outside of clicks. (D) Corresponding signal received on a single hydrophone, band-pass filtered between 500 and 2500 Hz.
ments. The synthetic aperture length between sensors is necessary to accumulate array gain. Subsequent analyses on the temporal and spectral characteristics of the clicks were performed on the noise-suppressed beamformed clicks.

2. Range estimation for sperm whales in the near- or far-field of the towed horizontal coherent receiver array

Each sperm whale individual was localized and tracked from its corresponding sequence of click bearing measurements using the moving array triangulation (MAT) technique (Gong, 2012; Gong et al., 2013), which combines bearing measurements from spatially separated apertures of a towed horizontal receiver array and employs the conventional triangulation ranging algorithm to localize a source in either the near- or far-field of the array. The whale range from the receiver array center was calculated using Eqs. (1) through (3) of Gong et al. (2013) given a pair of whale bearing measurements. The synthetic aperture length $A_s$ created by the array aperture.

3. Simultaneous depth and range estimation for a sperm whale in shallow water

The range and depth of a sperm whale located approximately 1 km from the receiver array in the Gulf of Maine were simultaneously estimated using MR-TDA of beamformed direct and singly or multiply reflected click signals from sea bottom and surface. The concept for whale range and depth inference is similar to that presented in Thode et al. (2002). However, because the array depth was accurately known from depth sensor measurement sampled every 10 ms, there was one fewer unknown. As a result, only three arrivals: Direct path, bottom bounce, and surface bounce were required to solve for whale range and depth as derived and discussed in the appendix. When more than three arrivals were present, the localization result could be obtained with higher accuracy by employing all available information (see the appendix). The whale range estimates obtained using MR-TDA will be compared to those obtained with bearings-only MAT method in Sec. IV.

B. Inferring sperm whale size from IPIs

The first 10-15 ms of a sperm whale click usually consists of multiple pulses a few milliseconds apart (Backus and Schevill, 1966; Mohl, 2001; Mohl et al., 2003; Norris and Harvey, 1972), resulting from multiple reflection within the whale head according to the bent-horn hypothesis (Norris and Harvey, 1972; Zimmer et al., 2003, 2004). The IPI has been shown to correlate with the spermaceti length (Gordon, 1990; Rhinelander and Dawson, 2004) and with the overall body size (Antunes et al., 2010; Growcett et al., 2011;
An automated moving local peak energy detector with a 1 ms averaging-time window was applied to the beamformed pressure-time series data to determine the arrival time of each pulse within a sperm whale click signal to estimate the IPI. The result for each click was plotted and visually inspected for accuracy. This analysis was applied to high-pass filtered beamformed pressure-time series data to suppress ambient noise and improve estimates of the IPI. Many click signals from each whale were examined, and only those with a clear multi-pulse structure were included in our analysis. IPI estimates were averaged over multiple clicks (roughly 20–60 per whale) to reduce the error in the IPI estimates. As can be noted from Table II, the standard deviation in the IPI estimates for each sperm whale is comparatively small, less than 10%. The whale body length, \( L_{\text{wh}} \), was then estimated here using the methods proposed by Gordon (1990) and Growcott et al. (2011)

\[
L_{\text{wh}, \text{Gordon}} = 4.833 + 1453 \times \text{IPI} - 0.001 \times \text{IPI}^2, \tag{1}
\]

\[
L_{\text{wh}, \text{Growcott}} = 1.257 \times \text{IPI} + 5.736. \tag{2}
\]

The sampling frequency of the towed array hydrophones limited our analysis of sperm whale click signal to frequencies \( \leq 2.5 \) kHz. In this low-frequency regime, the sperm whale multipulsed-click signals are approximately omnidirectional.

### C. Estimating sperm whale maximum detection range with the low-frequency towed coherent receiver array system

The maximum detection range of a sperm whale with our towed array system was determined as the range at which the transmission loss correction led to a received sperm whale click signal level that stood two standard deviations above the mean beamformed background noise level band-pass filtered between 300 Hz and 2.5 kHz. The two standard deviation signal excess enabled positive detection of sperm whales with over 95% confidence.

The broadband transmission loss from the sperm whale location to the receiver array was calculated at 10 Hz interval in the bandwidth of the received clicks using the parabolic equation-based range-dependent ocean waveguide-acoustic propagation model RAM (Collins, 1994). To simulate the effect of internal waves that randomized the acoustic propagation path, water-column sound speed profiles obtained from CTD and XBT measurements during the experiment were randomly updated every 500 m range, the approximate horizontal correlation length for linear internal waves. The bottom was assumed to be sandy with sound speed 1800 m/s, density 1800 kg/m\(^3\), and attenuation 0.8 dB/\( \lambda \). The water-column attenuation was set to \( 6 \times 10^{-5} \) dB/\( \lambda \). These waveguide parameters were previously found to provide the best-fit match between measured and modeled transmission loss in the Gulf of Maine (Andrews et al., 2009; Gong et al., 2010). The broadband transmission loss is obtained by incoherently averaging the received intensity from 50 Monte-Carlo realizations over the bandwidth of a sperm whale click from 1 to 2.5 kHz, following the approach of Andrews et al. (2009) and Gong et al. (2010).

### III. RESULTS I: A SPERM WHALE IN SHALLOW WATERS OF THE GULF OF MAINE

#### A. Analysis of vocalizations

Vocalizations from a sperm whale individual at site A in the shallow-waters of the Gulf of Maine were recorded using the towed receiver array for over an hour from 16:35:00 to 17:40:00 EDT on May 14. An example of the beamformed time series and spectrogram of an echolocation click train is shown in Fig. 3. These measurements were made in water-column depths of roughly 160 m where the receiver array was located at roughly 65 m depth. The average inter-click interval was approximately 0.7 s; this implied that these clicks could be categorized as “usual clicks” (Whitehead and Weilgart, 1990). Spectrogram analysis indicates that the clicks contain significant energy at low frequencies in the 1.5–2.5 kHz range. This enabled the beamformed, high-pass filtered, time-series data of the clicks to stand between 15 and 30 dB above the mean background ambient noise level [Fig. 3(C)]. Click rates were found to vary within an echolocation click train, as shown in Fig. 3, where the inter-click intervals and click amplitudes decrease slowly over a time interval of about 20 s. This implied that the sperm whale vocalizations could be transitioning from clicks to clicks, which are a sequence of low energy clicks closely spaced in time emitted when homing in on a prey (Madsen et al., 2002b; Miller et al., 2004).

Besides these usual echolocation clicks, we also recorded intense broadband clicks with dominant energy contained in the 0.5–2 kHz frequency range. These loud clicks were separated by intervals of 5 s or longer (see Fig. 4) and can be categorized as slow clicks (Barlow and Taylor, 2005; Jacquet et al., 2001; Oliveira et al., 2013; Weilgart and Whitehead, 1993). The spectra of both the echolocation and slow clicks recorded here closely match those of DTAG-recorded sperm whale click vocalizations shown in Fig. 1 of Oliveira et al. (2013).

The occurrences of the recorded vocalizations are shown against time in Fig. 7 with click types indicated. Each echolocation click train lasted roughly 2 min with periods of silence varying between 20 s and several minutes. Longer periods of silence lasting between 5 and 10 min observed here may be associated with the sperm whale’s ascent in the water-column (Zimmer et al., 2004) and resting near the surface.

By analyzing each received click in the time interval from 17:18:00 to 17:33:00 EDT, we estimated the mean IPI for this sperm whale individual to be 3.0 ms with a standard deviation of 0.3 ms. This corresponds to a sperm whale length of approximately 9.3 m, according to Eq. (2).

#### B. Tracking range and depth of a sperm whale close to array near-field

The estimated bearings of the vocalizations obtained via time-domain beamforming are plotted in Fig. 7. The instantaneous sperm whale ranges were estimated from the...
measured whale bearings using the bearings-only MAT technique for the time interval from 17:08:00 to 17:33:00 EDT and plotted in Fig. 8(A).

In the shallow water Gulf of Maine environment, clicks arriving at the receiver array from the sperm whale in close range to the array near-field distance display clear patterns of multiple reflection between the sea bottom and surface. The evolution of multipath arrival times is shown for all the clicks recorded between 17:18:00 and 17:33:00 EDT in Fig. 5. The direct arrival time of each click was first determined and aligned in time before the stacked sequence of clicks was created. Between 17:18:00 and 17:23:00 EDT, the first-order bottom and surface reflected arrivals are distinguishable in time, crossing each other due to vertical displacement of the whale. At other time instances, the bottom and surface reflections are not clearly distinguishable. The higher-order reflected arrivals are more prominent with shorter time separation for the later clicks in the time period analyzed. Multiple reflections in the shallow water waveguide extends the received sperm whale click time duration from 10-30 ms of the direct arrival (Møhl, 2001) to more than 250 ms.

The sperm whale range and depth were also simultaneously estimated by applying the MR-TDA technique described in the appendix for the time interval from 17:18:00 EDT to 17:33:00 EDT. All estimated delay times were used for each recorded click, giving a maximum of \( C_2 = 21 \) range-depth estimates when seven reflections were detected, and a minimum of \( C_2 = 6 \) range-depth estimates when four reflections were detected when employing the MR-TDA method. When ranges were available from the MAT technique, between four and seven depth-estimates could be obtained for each click, depending on the number of available reflections. The results are plotted in Fig. 8 comparing the MAT based results with the MR-TDA based results.

The analysis indicates that as the receiver array was being towed in the North bound direction, the sperm whale initially located roughly 0.5 km away from the array at 17:08:00 EDT moved away to a range of about 1.5 km in roughly 25 min. The whale appeared to hover within the water column without surfacing to breathe with estimated depth varying between 70 and 100 m over this time period. The increasing whale-receiver separation estimated using the MR-TDA explains the more compact arrival structure in Fig. 5 because when range increases, the reflection arrivals get closer to the direct path, as demonstrated in Fig. 12. The sperm whale range and depth estimates obtained via the MAT and MR-TDA methods agree well, especially near array broadside (between 17:23:00 and 17:28:00). At other time instances, the two methods yield results that are within 1 SD of each other. The whale range estimates obtained via MAT are expected to be more reliable than those obtained using MR-TDA. This is because in the MR-TDA technique, the whale range estimates depend on other parameters such as the unknown whale depth and water column depth. The range estimates obtained via MAT are not dependent on these parameters, instead it depends only on the measured bearing change of the whale with time.

C. Sperm whale detection range in shallow water

The click signals from this sperm whale located at roughly 1 km in range from the receiver array stood by as much as 35 dB over the mean background ambient noise level after beamforming [see Figs. 3(C) and 4(C)]. Because the background ambient noise level in the beam of the sperm whale after high pass filtering has roughly 5.5 dB standard.
deviation, the sperm whale click signals have a signal excess of \((35 - 2/5.5)^{1/2} \approx 24\) dB for our detector. The broadband transmission loss from a whale located approximately 1 km in range from the receiver array center and at a depth of 80 m is roughly 55 dB re 1 m [see Fig. 9(A)]. The signal excess of 24 dB implies that the sperm whale vocalizations can be detectable out to much longer ranges >60 km, where the transmission loss is \((55 + 24) = 79\) dB re 1 m, after beamforming with the towed receiver array. In contrast, sperm whale detection range with a single omnidirectional hydrophone is expected to be significantly limited [compare Fig. 4(C) showing the result after array beamforming with Fig. 4(D) which is the result for a single hydrophone]. Because sperm whale detection range is dependent on ambient noise level, the detection ranges given here are valid for low to moderate sea state of around two to three and wind speed between 8 and 11 kn, according to recorded measurements. The detection ranges will be larger at lower sea states and smaller at higher sea states.

IV. RESULTS II: MULTIPLE SPERM WHALES AT THE CONTINENTAL SLOPE SOUTH OF CAPE COD

We identified over 1000 sperm whale clicks in about 75 min of recording on May 13 at site B (Fig. 1) on the continental slope south of Cape Cod. Both left-right bearing estimates of the detected clicks are shown for roughly an hour of recording from 23:00:00 EDT on May 13 to 00:02:30 EDT on May 14 in Fig. 10. The left-right ambiguity of the linear receiver array is resolved for each group of clicks using the technique described in Sec. II A 1. The correlation coefficients between the change in bearings and the change in array headings for nine identified clusters of clicks in Fig. 10 are listed in Table I for both the left and right bearing candidates. A series of click bearings is determined to be true if (a) the correlation coefficient \(\rho\) is below 0.4, whereas the ambiguous group has \(\rho > 0.6\), or (b) the correlation coefficient \(\rho\) is at least five times smaller than that of the ambiguous group. When none of these criteria are met, the true click bearings could not be determined and no localization result is obtained.

A. Distinguishing sperm whale individuals using temporal, spectral, and spatial characteristics of clicks

We further associate the different click clusters to distinct sperm whale individuals based on the mean and standard deviation of IPI of each click cluster. The mean and standard deviation of the IPI as well as estimate of whale body length based on Eqs. (1) and (2) are provided in Table II for all

<table>
<thead>
<tr>
<th>Cluster number</th>
<th>(\langle IPI \rangle) (ms)</th>
<th>(\sigma_{IPI}) (ms)</th>
<th>(L_{w,Gordon}) (m)</th>
<th>(L_{w,Growthett}) (m)</th>
<th>Whale group</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.4</td>
<td>0.4</td>
<td>9.8</td>
<td>10.0</td>
<td>A</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>7.7</td>
<td>0.7</td>
<td>16.0</td>
<td>15.4</td>
<td>B</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>3.3</td>
<td>0.3</td>
<td>9.6</td>
<td>9.9</td>
<td>A</td>
<td>8</td>
</tr>
<tr>
<td>4</td>
<td>4.2</td>
<td>0.3</td>
<td>11.0</td>
<td>11.0</td>
<td>C</td>
<td>11</td>
</tr>
<tr>
<td>5</td>
<td>4.3</td>
<td>0.2</td>
<td>11.0</td>
<td>11.1</td>
<td>C</td>
<td>9</td>
</tr>
<tr>
<td>6</td>
<td>7.6</td>
<td>0.7</td>
<td>15.8</td>
<td>15.3</td>
<td>B</td>
<td>5</td>
</tr>
<tr>
<td>7</td>
<td>4.9</td>
<td>0.3</td>
<td>11.9</td>
<td>12.0</td>
<td>D</td>
<td>12</td>
</tr>
<tr>
<td>8-1</td>
<td>4.7</td>
<td>0.3</td>
<td>11.6</td>
<td>11.6</td>
<td>E</td>
<td>43</td>
</tr>
<tr>
<td>8-2</td>
<td>4.6</td>
<td>0.3</td>
<td>11.5</td>
<td>11.5</td>
<td>E</td>
<td>12</td>
</tr>
<tr>
<td>8-3</td>
<td>4.5</td>
<td>0.4</td>
<td>11.4</td>
<td>11.4</td>
<td>E</td>
<td>18</td>
</tr>
<tr>
<td>8-4</td>
<td>4.5</td>
<td>0.4</td>
<td>11.4</td>
<td>11.4</td>
<td>E</td>
<td>33</td>
</tr>
<tr>
<td>9</td>
<td>6.9</td>
<td>0.5</td>
<td>14.8</td>
<td>14.4</td>
<td>F</td>
<td>17</td>
</tr>
</tbody>
</table>
identified click clusters. Based on the measured IPI, we associate click clusters 1 and 3 (mean IPI \( \approx 3.4 \) ms) to the same whale A. Similarly, click clusters 2 and 6 are associated with whale B with much longer IPI of approximately 7.6 ms. Whale C is associated with click clusters 4 and 5. Click cluster 9 has very distinct IPI and corresponding whale size estimate and is associated with whale F. Whale D is associated with click cluster 7, while whale E is assumed to possess the IPI measurement of click cluster 8, which is composed of sub-clusters 8-1, 8-2, 8-3 and 8-4. The IPI values measured for clusters associated with whale E lie between those of whale C and whale D, so whale E could also possibly be either C or D. From Table II, the estimated number of sperm whale individuals simultaneously and passively recorded by the towed receiver array is at least 4 or as many as 6.

B. Localization and tracking of multiple sperm whales in far-field of the towed horizontal coherent receiver array

The sperm whale click clusters from May 13 are localized using the bearings-only MAT technique and the results are shown in Fig. 11 grouped according to the classification and association results obtained in Sec. IV A. Whale A

![Fig. 11](image)

![Fig. 12](image)

FIG. 11. (Color online) Localization and tracking result for multiple sperm whales with bearing clusters shown in Fig. 10 on May 13 at site B on the continental slope. The dashed curve is the track of the receiver array. The origin of the coordinate system is located at (39°50.94' N, 70°19.55' W).

FIG. 12. Calculated time delays from direct arrival for the bottom bounce \( \Delta t_b \) (upper left), surface bounce \( \Delta t_s \) (upper right), bottom-surface bounces \( \Delta t_{bs} \) (lower left), and surface-bottom bounces \( \Delta t_{sb} \) (lower right) as a function of whale range and depth. The receiver depth is set at 65 m and the water depth is 160 m.
(clusters 1 and 3) moves west a distance of about 1.1 km in a time interval of 17 min at a mean speed of roughly 1.1 m/s. Whale B (clusters 2 and 6) also moves west 2 km in 28 min, at a speed of approximately 1.2 m/s. Whale C (clusters 4 and 5) moves about 1.3 km in 10 min at a speed of 2 m/s. Whale D (cluster 7) is associated with a short period of clicking, and therefore no detailed movement tracking is available, although we can deduce that it is moving in the southwest direction. Whale E (cluster 8) has IPI values close to that of both whale C and D; however, the localization results show that this whale is further to the Northwest of the receiver, and several kilometers away from both whales C and D. Whale F has distinct spectral and temporal characteristics demonstrated by low ICI and high IPI and is localized to the southwest of the receiver toward the end of the track.

V. CONCLUSION

Sperm whales in the New England continental shelf and slope have been passively localized in both range and bearing, tracked, and classified using a single low frequency (<2500 Hz), densely sampled, towed horizontal coherent receiver array system. Coherent array gain, which increases sperm-whale-click-signal-to-background-noise ratio, was exploited to first determine the sperm whale bearing or horizontal direction by time-domain beamforming, and next estimate the sperm whale range or horizontal distance to the receiver array center by the moving array triangulation technique. Multiple concurrently vocalizing sperm whales were distinguished and classified based on their horizontal spatial location and the IPIs of their click signals, which is highly correlated to sperm whale body length. The vast majority of sperm whales were located in the far-field of the horizontal receiver array system. An advantage of bearings-only range localization with a densely sampled, towed horizontal coherent receiver array system is that no additional information about the environment, such as bathymetry or sound speed profile, is needed to estimate source range in the far-field of the array.

While studies have focused on examining the function of sperm whale clicks and their behavior during their descent in deep waters (Jacquet et al., 2001; Madsen et al., 2002b; Miller et al., 2004; Oliveira et al., 2013), their behaviors in shallow water are not well understood due to the fact that sperm whales are not common in this environment. Here the depth profile is estimated for a sperm whale in the shallow waters of the Gulf of Maine with 160 m water-column depth. This whale was located close to the near-field distance of the horizontal coherent receiver array system enabling its depth to be estimated from time-of-arrival-differences of the direct and the multiply reflected click signal. Over a roughly 15 min time interval analyzed in this paper, this sperm whale swam a horizontal distance of over 500 m, emitting both echolocation and slow clicks, while being entirely submerged and hovering close to the mid-water column.

The densely sampled, towed horizontal coherent receiver array system employed here is similar to that used in naval operations for long range ocean surveillance and in geophysical exploration. It may be important and necessary to remotely sense marine mammal activity from long ranges in these applications. The methods developed here and the results obtained here can be directly applied to assess the performance of these other coherent receiver array systems for monitoring sperm whales from long ranges.

ACKNOWLEDGMENTS

This work was funded by the National Science Foundation and the Office of Naval Research. The authors would like to thank Scott Whitney of BAE Systems, Carl Stevens of Einhorn Engineering, and the staff and crew of RV Endeavor for assistance with the data collection.

APPENDIX: SIMULTANEOUS RANGE-DEPTH LOCALIZATION USING MR-TDA

A formulation and an analytic solution are first provided for simultaneous range-depth localization of a sperm whale using delay times of the two first-order reflections of its click from the sea surface and bottom. Next, a method incorporating all detected reflections, including higher-order reflections from sea surface and bottom is presented. This latter method can enhance localization accuracy and reduce errors associated with estimating click reflection delay times.

Let the known receiver array center depth be \( z_r \), and the unknown whale range and depth be \( r_w \) and \( z_w \), respectively. Let \( H \) be the water depth that is assumed to be constant within the propagation path from the whale to the receiver. The sound speed variation with depth and range is assumed to cause negligible refraction given the short propagation path from the sperm whale to receiver. The direct, first-, and second-order reflected path lengths from the vocalizing whale to the receiver are

\[
\text{Direct path: } d = \sqrt{r_w^2 + (z_r - z_w)^2}, \quad (A1)
\]

\[
\text{Surface-reflected: } s = \sqrt{r_w^2 + (z_r + z_w)^2}, \quad (A2)
\]

\[
\text{Bottom-reflected: } b = \sqrt{r_w^2 + (2H - z_r - z_w)^2}, \quad (A3)
\]

\[
\text{Surface-bottom reflected: } sb = \sqrt{r_w^2 + (2H + z_w - z_r)^2}, \quad (A4)
\]

\[
\text{Bottom-surface reflected: } bs = \sqrt{r_w^2 + (2H - z_w - z_r)^2}. \quad (A5)
\]

Because there are only two unknowns \( r_w \) and \( z_w \), only two time delays are needed to estimate these unknowns. We choose the two first-order surface-reflected and bottom-reflected arrivals as they often encounter the least losses and are therefore easier to detect. The difference in arrival times are

\[
\Delta t_s = \frac{s - d}{c} = \frac{4z_r}{c(2d + \Delta t_s)}, \quad (A6)
\]


Tran et al.: Sperm whale range localization and classification
\[ \Delta t_b = \frac{b - d}{c} = \frac{4(H - z_r)(H - z_w)}{c(2d + \Delta t_b)}, \tag{A7} \]

which can be combined into
\[ \begin{bmatrix} -2c\Delta t_s \\ -2c\Delta t_b \end{bmatrix} \begin{bmatrix} 4z_r \\ -4(H - z_r) \end{bmatrix} \begin{bmatrix} d \\ z_w \end{bmatrix} = c(\Delta t_s)^2 \begin{bmatrix} c(\Delta t_b)^2 \\ -4(H - z_r)H \end{bmatrix}, \tag{A8} \]

where \( c \) is the mean sound speed in the water-column. Using the time delay of the first-order surface and bottom reflected arrivals, \( \Delta t_s \) and \( \Delta t_b \), as inputs, the range and depth of the whale can be calculated via
\[ \begin{bmatrix} d \\ z_w \end{bmatrix} = \begin{bmatrix} -2c\Delta t_s \\ -2c\Delta t_b \end{bmatrix}^{-1} \begin{bmatrix} 4z_r \\ -4(H - z_r) \end{bmatrix} \begin{bmatrix} d \\ z_w \end{bmatrix} \times \begin{bmatrix} c(\Delta t_s)^2 \\ c(\Delta t_b)^2 \end{bmatrix}, \tag{A9} \]

\[ r_w = \sqrt{d^2 - (z_r - z_w)^2}. \tag{A10} \]

When more than three arrivals are present, the time delays for each order of multiple reflection can be calculated as a function of candidate source ranges and source depths given the water depth and receiver depth as inputs. Suppose a time delay \( \Delta t_0 = \tau_0 \) is measured for the first bottom reflection. A unique contour then exists for the set of range and depth combination \( \{r_w, z_w\} \) such that \( \Delta t_b = \tau_0 \). This is illustrated in Fig. 12 for various delays between 0 and 80 ms. Similarly, constant time delay contours can be constructed for the surface bounce, bottom-surface bounces, surface-bottom bounces (see Fig. 12), and other higher order multiple reflections. If two time delay measurements are available, an estimate of the whale location can be obtained at the intersection of the two contours. When the two time delays are the first-order reflections from ocean bottom and surface, this is equivalent to solving Eq. (A9). When \( N \) time-delays are available, the number of contour intersections is
\[ \binom{N}{2} = N(N - 1)/2, \]

which corresponds to the number of independent range-depth estimates based on time delay pairs. The solution is then taken to be the mean range and depth obtained from averaging these independent estimates.

In the shallow Gulf of Maine environment \( (H < 200 \text{ m}) \), multiple reflections from the sea bottom and surface are only distinguishable from the direct arrival at small ranges approximately within a couple of kilometers. At longer ranges, the time difference of arrival of the direct and first-order interface reflections \( \Delta t_b \) and \( \Delta t_s \) becomes smaller and negligible with increasing range as shown in Fig. 12. When the whale is far away from the receiver, only higher order reflections are separable in time from the direct path; however, they suffer higher transmission losses due to multiple bottom and surface interaction and are difficult to detect (see Fig. 1 of Thode et al., 2002). The small range also guarantees that the water depth \( H \) is relatively unchanged along the acoustic propagation path from whale to receiver.

1. Depth-localization when range information is available

When the range of a sperm whale can be estimated using an independent method, such as the MAT, the MR-TDA problem has reduced dimensionality because only the whale depth needs to be estimated. Figure 12 shows that a unique depth solution exists for each click reflection delay time measurement at a particular range. The whale depth is then estimated as the average of the \( N \) independent depth estimates from \( N \) reflected click time delays.


