Methods for tracking multiple marine mammals with wide-baseline passive acoustic arrays

Eva-Marie Nosal[a)]
Department of Ocean and Resources Engineering, School of Ocean and Earth Science and Technology, University of Hawaii at Manoa, 2540 Dole Street Holmes Hall 402, Honolulu, Hawaii 96822

(Received 28 June 2012; accepted 28 September 2012)

Most methods used to track marine mammals with passive acoustics require that time differences of arrivals (TDOAs) are established and are associated between hydrophone pairs. Consequently, multiple animal trackers commonly apply single-animal TDOA localization methods after performing a call separation and/or TDOA association step. When a wide-baseline array is used with multiple animals that make similar calls with short inter-call-intervals, the separation/association step can be challenging and potentially rejects valid TDOAs. This paper extends a model-based TDOA method to deal with multiple-animal datasets in a way that does not require a TDOA association step; animals are separated based on position. Advantageously, false TDOAs (e.g., a direct path associated with a multipath arrival) do not need to be removed. An analogous development is also presented for a model-based time of arrival tracking method. Results from simulations and application to a multiple sperm whale dataset are used to illustrate the multiple-animal methods. Although computationally more demanding than most track-after-association methods because separation is performed in a higher-dimensional space, the methods are computationally tractable and represent a useful new tool in the suite of options available for tracking multiple animals with passive acoustics.
© *2013 Acoustical Society of America*. [http://dx.doi.org/10.1121/1.4816549]

PACS number(s): 43.30.Sf, 43.30.Wi, 43.60.Jn [AMT] Pages: 2383–2392

I. INTRODUCTION

Passive acoustic monitoring (PAM) methods use signals recorded on hydrophone(s) to detect, classify, and track vocalizing marine mammals. PAM is valuable for long-term monitoring, behavioral studies, and bioacoustics research on free-ranging animals. PAM methods complement visual monitoring since calling animals can be detected underwater over long ranges and in conditions that are not favorable for visual monitoring (such as nighttime and rough seas). PAM methods are also useful during tagging studies since they can provide information about the movement of animals underwater, far-field bioacoustic signal properties, and other acoustic events in the area (including from other marine mammals).

A recent emphasis on improving PAM methods stems from concern over anthropogenic impacts on the marine environment, and considerable progress has been made in the field over the past several years as a consequence. Nevertheless, significant challenges persist. This paper is concerned with localization and tracking of marine mammals using an array of wide-baseline hydrophones. Numerous methods have been developed for this problem,[1–8] most of which make use of the time-difference-of-arrival (TDOA) of a call (or call sequence) on pairs of hydrophones. Most applications are limited to cases with a single animal or cases where calls from individual animals can be separated and/or a call from an animal can be associated over the hydrophones. Two general and common approaches to multiple-animal tracking are summarized below: (1) source separation followed by association and (2) source association without separation (or joint association/separation). Both of these approaches produce TDOAs that are associated with individual animals, and localization is performed using single-animal tracking techniques.

In the source separation approach, calls from individual animals are first separated on single hydrophone recordings and then associated between hydrophones. As an illustrative (albeit over-simplified) example, consider a mixed species dataset in which only one individual from each species is present and the different species have unique call characteristics; then calls can be attributed to individual animals. Such separation based on call characteristics is sometimes possible for single-species datasets if calls are sufficiently varied. Assigning calls to individuals may be possible, for example, with two humpback whales in different parts of a song[9,10] that may have very different source characteristics (at least for some limited time period), or with male bearded seal vocalizations that have clear individual variation.[11] Even with stereotyped impulsive (click-type) vocalizations, it can be possible to separate click trains from individual animals, for example by exploiting the unique and slowly varying multipath arrival structure for individual animals,[12] or by tracking slowly varying features of clicks within a train, such as inter-click-interval (ICI), power, or waveform.[13] Looking forward, advances in passive acoustic monitoring methods for detection and classification contribute new and useful techniques for source separation[14–16] that will aid in localization efforts as well. Once separated, call sequences can be associated between hydrophone pairs in various ways. This step would be automatic in our over-

a)Electronic mail: nosal@hawaii.edu

Redistribution subject to ASA license or copyright; see http://acousticalsociety.org/content/terms. Download to IP: 128.171.57.189 On: Wed, 16 Oct 2013 20:25:08

simplified mixed species example while in other cases it can be more difficult but is often possible. With clicking animals, for example, a "rhythm analysis" that finds a best match of the sequence of inter-call-intervals on one phone to those on the second phone can be useful.[17,18]

The association without separation (or joint association/separation) approach is useful in cases where is not possible or desirable to separate sources on individual hydrophones but it is possible to associate calls/call-sequences between receivers for TDOAs extraction. This approach can be relatively straightforward in cases where calling rates are low, such as with minke whale boings or blue whale calls; unless two animals happen to call at almost the same time, only one call will be received within the domain of physically realistic TDOAs (which is limited by the spacing of the hydrophones). Another example of this approach forms a cross-correlogram (cross-correlations of windowed and pre-conditioned signal waveforms or spectrograms at different time steps) in which slowly varying peaks correspond to persistent TDOA estimates and hence, potentially, to individual animals.[5,6,19–22] Confounding, incorrect associations between different multipath arrivals can also produce persistent peaks in cross-correlograms which can lead to "false TDOAs"; various methods for dealing with this problem have been explored in the literature.[23–25] Once TDOAs have been established, they must be associated between receiver pairs before single animal tracking methods can be applied. One approach is to impose the self-consistency relationship between TDOAs of hydrophone triplets,[23] i.e., $\Delta t_{ij} = \Delta t_{ik} - \Delta t_{kj}$ (where Δt_{ij} denotes the difference in arrival time of a call between phones i and j). Relying on TDOA triplets potentially rejects valid TDOAs (e.g., when calls are not detected on enough hydrophones and/or some TDOAs are missed in the TDOA estimation step), but in many cases this is a sufficient approach that is especially useful because it is simple and robust.

This paper considers a third approach to the multiple-animal tracking problem in which individual animals are separated and tracked based on location estimates. It may be helpful to think of this as a wide-baseline analogy to separating animals recorded on a short-baseline array by tracking slowly varying bearings.[17,26–30] The approach does not require a source-separation step, nor does it require that TDOAs are associated between pairs. Moreover, there is no need for multipath removal or any classification step to remove false or mis-associated TDOAs. The method builds on the so-called "model-based TDOA" framework for localization,[31–33] which seeks to minimize a weighted least-squares difference between measured and modeled TDOAs. A related approach was recently proposed by Baggenstoss,[18] who also handles the TDOA association and source tracking simultaneously. In his approach, position estimates are initialized from one set of pre-associated TDOAs, after which tracking is performed recursively through maximization of a log-likelihood function for position given a collection of (unassociated) TDOAs. The multiple-animal TDOA localization method proposed here uses a different likelihood function than the one used by Baggenstoss and does not require an initialization step or recursive application. However the output here is a list of location estimates (instead of tracks) which must be post-processed to form tracks. In addition to the multiple-animal model-based TDOA localization method, this paper also develops an analogous multiple-animal model-based localization method based on time-of-arrivals (TOAs) instead of TDOAs. This method is useful in cases where TOAs are available but may be difficult to associate to individual animals or between receiver pairs.

It has the added advantage of inverting for call time (in addition to position) which reduces position uncertainty by allowing some error to be attributed to call time. To investigate and demonstrate applicability, the model-based multiple-animal TDOA and TOA localization methods are applied to the same multiple-sperm whale dataset that has been analyzed by other researchers.[18,34]

II. BACKGROUND

This section introduces the notation and definitions used in the paper and briefly reviews the model-based TDOA and TOA methods.

A. Notation and definitions

An animal at position $\mathbf{w}$ emits a call at time t. The call propagates through the water to hydrophone i giving a time of arrival (TOA) denoted as t_i,

$$t_i = t_i(t, \mathbf{w}) = t + T_i(\mathbf{w}), \tag{1}$$

where $T_i(\mathbf{w})$ is the propagation time between $\mathbf{w}$ and hydrophone i. The difference in arrival time (TDOA) between two hydrophones is denoted by

$$\Delta t_{ij}(\mathbf{w}) = t_i - t_j. \tag{2}$$

Modeled estimates of travel times, TOAs, and TDOAs are functions of call time and/or position and are denoted with carets, e.g., $\hat{T}_i(\mathbf{w})$, $\hat{t}_i(t, \mathbf{w})$, and $\Delta\hat{t}_{ij}(\mathbf{w})$. Modeled TDOAs are calculated as

$$\Delta\hat{t}_{ij}(\mathbf{w}) = \hat{T}_i(\mathbf{w}) - \hat{T}_j(\mathbf{w}), \tag{3}$$

where the $\hat{T}_i(\mathbf{w})$ are times between $\mathbf{w}$ and hydrophone i calculated using some assumed or measured sound speed profile (SSP) and sound propagation model (e.g., a constant SSP with a straight-line propagation model or a depth-dependent SSP with a ray-tracing model).

The above applies to direct-path arrivals. In the presence of multipath arrivals (e.g., arrivals that bounce off the surface and/or bottom) the method of images can be used to represent the multipath contribution as a direct arrival received on an image receiver[4,7,33,35–38] located above the surface or below the seafloor. In all methods discussed below, if multipath arrivals are present then image receivers can be added to the collection of receivers and treated in the same manner as real receivers (thereby increasing the number of receivers used for localization).

B. Single animal TDOA method

Likelihood functions are formed as a weighted mean squared error between the measured and modeled time differences[18,25,31,33]

Redistribution subject to ASA license or copyright; see http://acousticalsociety.org/content/terms. Download to IP: 128.171.57.189 On: Wed, 16 Oct 2013 20:25:08

$$LS(\mathbf{w}) \propto \prod_{i,j} \exp\left[\frac{-1}{2\sigma_{ij}^2}\left(\Delta t_{ij} - \Delta \hat{t}_{ij}(\mathbf{w})\right)^2\right], \tag{4}$$

where σ_{ij} is the standard deviation in TDOA error due to uncertainties in receiver position, measured click times, and SSP. This surface assumes errors are normally distributed with the same distribution for each receiver pair and call. It also treats receiver position and sound speed as fixed instead of treating them as nuisance parameters. These simplifications are used to minimize computational burden that could be relaxed in a more rigorous treatment.[39–41]

To localize a call, the TDOA likelihood function is maximized over possible source locations $\mathbf{w}$ to find the best estimate of source location for a list of measured TDOAs. For a source heard on all receiver pairs, the value of the likelihood function approaches 1 at the correct source location provided that the TDOAs have been correctly measured and the SSP and sound propagation models are sufficiently accurate.

C. Single animal TOA method

In cases where it is possible to detect calls and identify precisely the same point in the call (e.g., possible with frequency modulated calls and short-duration clicks, but not with tonal calls), another useful approach is to match time of arrivals (TOAs) instead of TDOAs. This expands the search to call time as well as position. That is, we maximize the following likelihood surface[32] over all locations and time $t < t_i$ (by causality; the signal cannot arrive before it was produced)

$$\begin{aligned} LS(t, \mathbf{w}) &\propto \prod_i \exp\left[\frac{-1}{2\sigma_i^2}\left(t_i - \hat{t}_i(t, \mathbf{w})\right)^2\right] \\ &= \prod_i \exp\left[\frac{-1}{2\sigma_i^2}\left(t_i - \hat{T}_i(\mathbf{w}) - t\right)^2\right], \end{aligned} \tag{5}$$

where σ_i is the standard deviation in the TOA errors due to uncertainties in receiver position, measured click times, and SSP. This surface incorporates the same simplifications that are used in the TDOA surface to reduce computational burden.

Using TOA instead of TDOA can be advantageous since call times are estimated directly (treated as parameters) in addition to call positions which reduces error in position estimates by allowing some errors to be "absorbed" in the call time estimate instead of only in position estimate (as is the case with TDOA methods). Moreover, the TOA method does not require TDOAs to be established, which can be useful in datasets with multiple animals and/or with reflections since establishing TDOAs can be difficult in these situations.

III. METHODS

This section develops two approaches to multiple animal tracking. The approaches are generalizations of the single-animal model-based TDOA and TOA methods.

A. Multiple animal TDOA method

The generalized TDOA method assumes that TDOAs for each hydrophone pair have been established. It does not require that TDOAs have been associated between receiver pairs; the TDOA on one receiver pair does not need to be associated with the TDOA of another receiver pair. Furthermore, it does not require that all TDOAs are valid; all potential TDOAs are used with no effort made to separate TDOAs for individual animals or to prune "false" TDOAs (for example, the result of a multipath arrival being paired with a direct arrival, or of a call from one animal being paired with a call from another animal). That is, for each hydrophone pair, a list of all potential TDOAs is input to the localizer with no other information about the classification or association of these TDOAs.

For a given time step, the multiple-animal TDOA surface incorporates all TDOAs for each hydrophone pair as

$$LS(\mathbf{w}) \propto \prod_{i,j}\left\{\max_k\left(\exp\left[\frac{-1}{2\sigma_{ij}^2}\left(\Delta t_{ij}(k) - \Delta \hat{t}_{ij}(\mathbf{w})\right)^2\right]\right)\right\}, \tag{6}$$

where $\Delta t_{ij}(k)$ is the kth measured TDOA that falls within the given time step for receiver pair $\{i, j\}$. The TDOA surface attains a local maximum with value close to 1 for a true source heard on all hydrophone pairs considered. It may also attain a local maximum near 1 at a non-source location due to spurious agreements between measured and modeled arrivals (corresponding to incorrect or mis-associated TDOAs). However, if enough hydrophones are used, peaks that correspond to a true source will persist over time (from one call to the next) while incorrect/spurious peaks will not persist. The incorrect/spurious peaks can be eliminated in a post-processing step that clusters persistent location estimates into a track.

Intuitive understanding of how this processor works can be gleaned from Fig. 1 which shows the contribution to the surface from one receiver pair in a simulated dataset with two animals in two dimensions. Each established TDOA corresponds to one band of high likelihood (black). This figure also demonstrates the reason for using maxima over all TDOAs instead of sums in Eq. (6); if sums were used, two closely separated TDOAs would add to form one peak centered between the peaks for each individual TDOA instead of two overlapping peaks, as desired. When this surface is multiplied with a surfaces from another hydrophone pair, only the positions that agree are reinforced. Multiplication over sufficiently many receiver pairs eliminates all but the correct sources. Incorrect positions potentially remain even after using all hydrophone pairs. However, if enough hydrophones are used, these incorrect positions do not persist over time while the correct positions do. Although the illustration is for two dimensions, the idea in three dimensions is similar: the likelihood subsurface for each receiver pair has multiple high likelihood value surfaces (hyperboloids for constant SSP and isodiochrons[42] for non-constant SSPs) which agree over multiple receiver pairs at the potential source positions.

Redistribution subject to ASA license or copyright; see http://acousticalsociety.org/content/terms. Download to IP: 128.171.57.189 On: Wed, 16 Oct 2013 20:25:08

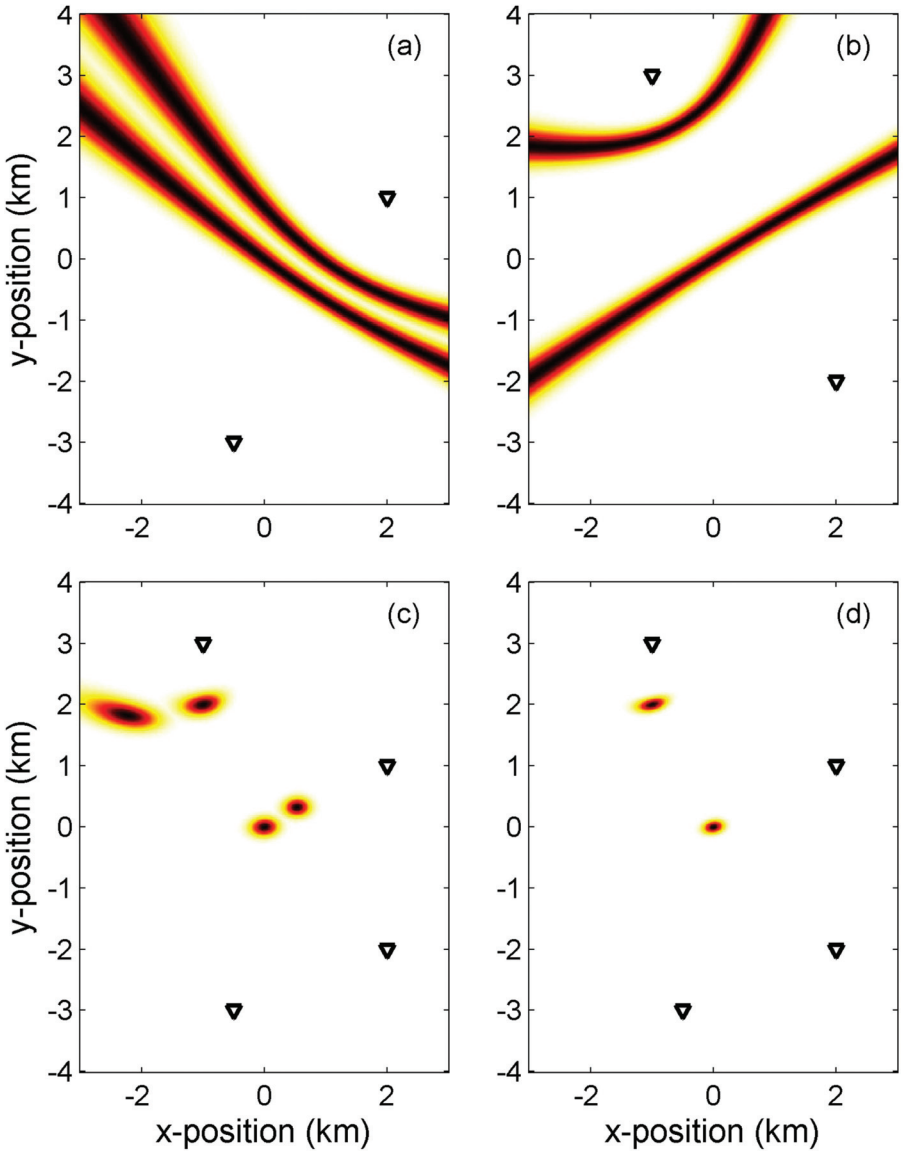


FIG. 1. TDOA likelihood surfaces for a simulated 2D, two animal dataset corresponding to (a) two TDOAs heard on hydrophones shown with triangles; (b) two TDOAs heard on a different pair of hydrophones shown with triangles; (c) entry-wise multiplication of the sub-surface from (a) and (b); (d) the use of all hydrophone pairs. The two peaks in (d) correspond to the position of the two simulated animals. Black/red/white indicate high/medium/low likelihood values, respectively.

B. Multiple animal TOA method

For the generalized TOA method, we assume that TOAs for each hydrophone have been established but that the TOAs have not been classified (i.e., labeled as direct or multipath arrivals), separated (i.e., assigned to different individuals), or associated (i.e., established which TOA on one hydrophone corresponds to the TOA of another hydrophone). In summary, for each hydrophone, we have a list of all detected TOAs but no other information.

The multiple-animal TOA likelihood surface incorporates all TOAs for each hydrophone as

$$LS(\mathbf{w},t) \propto \prod_i \left\{ \max_k \left(\exp\left[\frac{-1}{2\sigma_i^2} (t_i(k) - \hat{T}_i(\mathbf{w}) - t)^2 \right] \right) \right\}, \tag{7}$$

where $t_i(k)$ is the kth measured TOA on hydrophone i such that $t_i(k) > t$ (by causality). The TOA surface obtains a maxima near 1 for a correct source location and time. As with the TDOA method, a maxima near 1 might be attained for an incorrect/spurious source, but such false maxima will not persist over time and hence can be eliminated in a post-processing step for animal tracks. Figure 2 illustrates this processor using the same simulated dataset used in Fig. 1. Surfaces at different times have circular maxima at different radii from the receiver locations; as the search time increases, the circles shrink down to the receiver position (corresponding to a click time equal to the search time) and then disappear. Generalizing to three dimensions gives a family of spheres (or TOA isochrons for non-constant SSP) for each receiver.

C. Required number of hydrophones

The multiple-animal TDOA and TOA methods require four and five hydrophones to localize unambiguously in two and three dimensions, respectively. The reason for these requirements are the same as those for the single animal TDOA methods, for which careful discussion can be found in the literature.[43,44] The minimum number of required phones can be reduced for certain hydrophone geometries, by using physical constraints (e.g., seafloor, sea-surface),[35,43] and/or by using of multipath arrivals.[4,7,33,36]

Redistribution subject to ASA license or copyright; see http://acousticalsociety.org/content/terms. Download to IP: 128.171.57.189 On: Wed, 16 Oct 2013 20:25:08

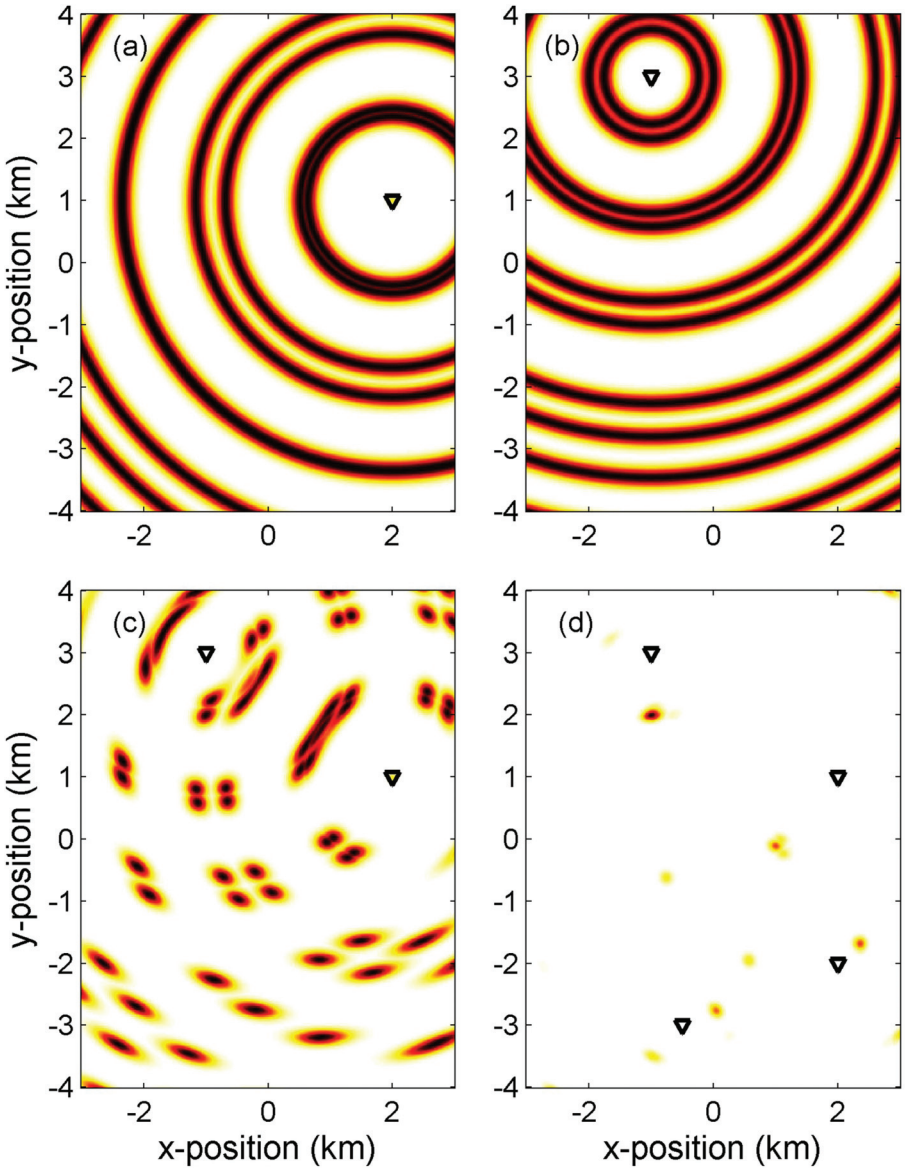


FIG. 2. TOA likelihood surfaces at a single time for the simulated dataset corresponding to (a) and (b) multiple TOAs detected on the hydrophone indicated with a triangle; (c) entry-wise multiplication of the sub-surfaces from (a) and (b); (d) the use of all hydrophones. Only one simulated animal is found, corresponding to a call that was emitted by the animal at the search time. Several lower-amplitude spurious local peaks are also evident. The second animal appears when the surface is recalculated at a time corresponding to a click emitted by the second animal.

For the multiple-animal TDOA and TOA methods, products taken over all hydrophones are maximized near a value of 1 at the correct source time/location only if the call is detected on all hydrophone pairs (hydrophones). If a TDOA (TOA) is missed on one or more hydrophone pairs (hydrophones), the likelihood value in the factor corresponding to the missed TDOA (TOA) will be near zero so that multiplication with the other factors will yield a value near zero. Correspondingly, if there are more hydrophones available than the minimum required for unambiguous localization and if it is possible that calls are detected on only a subset of the hydrophones, it might be desirable/necessary to maximize over several versions of the full likelihood surface; one corresponding to each subset of hydrophones to be considered (where each subset contains at least the minimum number of required phones). In these cases Eqs. (6) and (7) can be modified to limit the product to the "best M hydrophones" (where M is the number of phones to be used for localization); that is, only the M hydrophones that give the maximum likelihood value for a given search point are used in the product and the others are discarded. This is a simple process to implement in practice since the factors inside the product can be calculated and sorted, with only the factors corresponding to the best M hydrophones used in the product.

D. Finding peaks and establishing tracks

In the multiple-animal TDOA and TOA methods, tracking is performed by finding persistent (i.e., slowly moving in position over time) peaks in the resulting likelihood surfaces. This is not a trivial problem, but is simpler for the TDOA method for two main reasons: (1) TDOA has a smaller parameter space since it requires a search over space while TOA requires a search over space and time and (2) TDOA has fewer false peaks because the process of establishing TDOAs reduces the number of associations, while the TOA method retains all possible associations. In the author's experience, a simple grid search (with sufficiently small grid spacing) to find maxima (i.e., localizations) together with a clustering technique to group localizations into tracks is often adequate for the TDOA method. The clustering method

Redistribution subject to ASA license or copyright; see http://acousticalsociety.org/content/terms. Download to IP: 128.171.57.189 On: Wed, 16 Oct 2013 20:25:08

should include practical limitations, such as those on swim speed and minimum track durations. Instead of a grid search, a simple optimization scheme was used in the TDOA example below, which worked efficiently and well. Unfortunately, the grid search and subsequent clustering approach becomes computationally prohibitive for the TOA method because of the added time dimension in the search space. More sophisticated search algorithms are needed to overcome this difficulty, with only one of many possible approaches outlined in the example below.

IV. APPLICATION TO DATA

A. Sperm whale dataset

The datasets investigated here was provided by the Naval Undersea Warfare Center for the Second International Workshop on Detection and Localization of Marine Mammals using Passive Acoustics. Data were recorded at the Atlantic Undersea Test and Evaluation Center (AUTEC) located in the Tongue of the Ocean off Andros Island, Bahamas. One dataset contains 25 min from a single clicking sperm whale recorded on five hydrophones and another dataset contains 20 min from of multiple clicking sperm whales recorded on six hydrophones, although excessive noise on one phone meant only five phones could be used for tracking. The hydrophones were bottom-mounted at a depth of about 2 km and spaced by several kilometers. The sampling rate was 48 kHz. Further details (including hydrophone positions) can be found in Adam *et al.*[45] The single animal dataset has been extensively analyzed in other work.[18,23,25,32,38,46] The multiple animal dataset, which is the focus of the effort presented in this section, has also been previously examined by several other researchers[13,34] which provide a useful comparison to the results obtained here.

B. Click detection

Clicks were detected using an automated transient detector based on Page's test[35] applied to the envelope of the data after high-pass filtering using a 300 Hz high-pass Butterworth filter. Complete details on the implementation can be found elsewhere.[32] All clicks detected were used in the following analysis, with no attempt to classify clicks as direct or surface-reflected.

C. TDOA

TDOAs were extracted using scatterplots of TDOA vs time. For a given hydrophone pair, all possible click associations and corresponding TDOAs were formed by pairing all clicks on the first phone with all clicks on the second phone that gave physically viable TDOAs; the magnitude of direct-arrival TDOAs must be less than the travel time between the two phones, corresponding to a maximum of approximately 5 s for the dataset used here. When plotted as a function of the arrival time on one of the hydrophones, potentially valid TDOAs appear as slowly varying lines and mis-associations appear as random scatter (Fig. 3).

Slowly varying tracks were extracted from the scatterplots of TDOA vs click time using an automated algorithm.

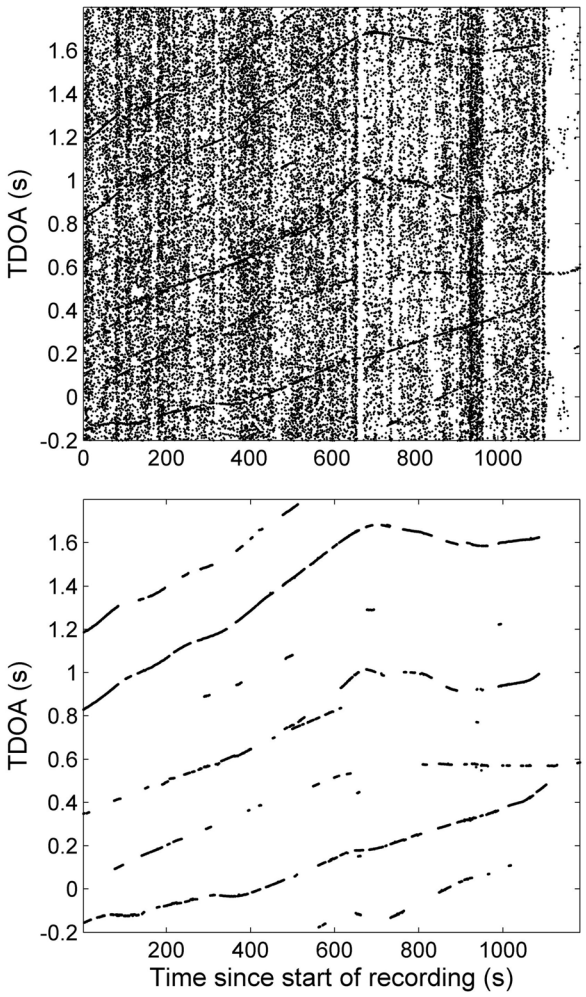


FIG. 3. TDOA vs click time scatterplot for two phones from the multiple-animal dataset (top) and TDOAs extracted using the TDOA tracking method described in Sec. IV C (bottom).

Since each track corresponds to a click train from a single animal, known features of sperm whale click trains could be used to extract the trains. The following conditions were used:

(1) ICIs must be within the range known for sperm whale usual clicks ($0.2\,\text{s} \leq \text{ICI} \leq 2\,\text{s}$).
(2) ICIs vary slowly. This was quantified by the following condition: $|\text{ICI}_i - \text{ICI}_{i+1}|/\text{ICI}_i < 0.8$ where ICI_{i+1} directly follows ICI_i in an extracted TDOA sequence.
(3) The rate of change of TDOA is limited by the maximum swim speed of a sperm whale. The maximum TDOA change rate will occur for an animal swimming along a straight line joining the two hydrophones at maximum speed. Using a sperm whale maximum swim speed of about 10 km/h (highly overestimated to be conservative) and a nominal sound speed of 1500 m/s, this corresponds to a maximum TDOA change of 0.002 s/s. This defines a maximum possible slope magnitude on the TDOA vs click time scatterplot.

For every point on the scatterplot, all secondary points that satisfied conditions (1) and (3) above were considered as a candidate second point for a track beginning at the first

Redistribution subject to ASA license or copyright; see http://acousticalsociety.org/content/terms. Download to IP: 128.171.57.189 On: Wed, 16 Oct 2013 20:25:08

point. For a given candidate second point, a line joining points one and two was formed. The point satisfying all three conditions that fell closest to the line joining the first and second points was chosen as a third point. If no points satisfied the conditions the track was terminated. This process was iteratively repeated using the last two points in the track at each step to choose the next best point. A track was retained if it contained at least 10 points. Figure 3 shows an example of the resulting scatterplot and TDOA tracks.

Not all slowly varying lines extracted from the TDOA vs click time scatterplots correspond to correct direct-arrival associations (e.g., Fig. 4). Persistent associations between one arrival path (e.g., direct) on one phone with those of another arrival path (e.g., surface-reflected) from the same animals on the second phone will also appear as slowly varying lines; these correspond to "false" TDOAs. Multipath arrivals on the first phone can also be associated correctly with the corresponding multipath arrivals on the second phone, potentially providing additional information on animal position if multipath arrivals are used in the localization via the virtual receiver approach.

One of the main benefits of using the multiple-animal TDOA method is that it is not necessary to classify the extracted TDOAs as direct-direct, reflected-reflected, or mis-associated. Since Eq. (6) takes the TDOA that gives the maximum contribution to the likelihood surface, if the correctly associated TDOA is available then it will be used instead of mis-associations when calculating the likelihood value at the correct animal location. The mis-associations will potentially contribute to incorrect peaks in the likelihood surface, but it is unlikely that all receiver pairs will agree on the same spurious peak and if they do, the agreement will not persist. In addition, there is no need to further associate TDOAs (e.g., association of TDOA for receiver pair {1,2} with a TDOA on receiver pair {1,3}) as would be necessary if a single-animal tracking method were being used.

Once potential TDOAs were established for each receiver pair, the multiple-animal TDOA method was used to find the animal positions. For this dataset the estimated contribution to σ_i due to uncertainty in (1) measured click time was 5 ms, (2) receiver position was 2 ms, and (3) SSP was 3 ms (more details on the assignment of these values can be found in Nosal and Frazer[32]). This gives $\sigma_i = \sqrt{5^2 + 2^2 + 3^2} \approx 6.2$ ms. TDOA is calculated as the difference in TOAs which gives $\sigma_{ij} = \sqrt{\sigma_i^2 + \sigma_j^2} = \sqrt{2\sigma_i^2} \approx 8.7$ ms. Time steps of 2.5 s were used, with all measured TDOAs within a 5 s window centered on a given time step input into Eq. (6). In the first pass, all five hydrophones were used. A constrained Nelder–Mead multidimensional optimizer initialized with a point in the center of the hydrophone array was used to find one maximum in the resulting likelihood surface. The optimizer was constrained to limit the search to within the water column. The same speed profile as used in Nosal and Frazer[38] was used to create a look-up table of travel time vs range (100 m grid steps) and depth (25 m grid spacing) from which the modeled travel times were interpolated during the search. The TDOAs that corresponded to the maximum position in the likelihood surface were eliminated from the list of TDOAs and the minimizer was applied again. This process was repeated iteratively, finding maxima and eliminating TDOAs, until no maxima above a certain threshold (set here at 0.95) remained for the current time step. This gave a list of potential source positions for clicks heard on all five phones. A second pass was run on the remaining TDOAs using four phones to localize; since the hydrophones were on the seafloor, ambiguity could be reduced by requiring the sources to be within the water column, thus reducing the number of required phones from 5 to 4. Figure 5 shows the resulting position estimates. This iterative technique was used instead of a global optimization method since it proved to be computationally more efficient.

Uncertainty in position estimates was quantified as the 90% central credible interval defined here as the interval that leaves out 5% area on both tails of the conditional distribution for the parameter of interest (x, y, or z) at the estimated animal position (although not ideal, conditionals were used

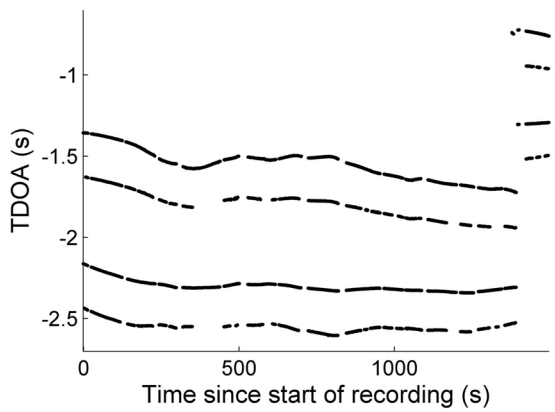


FIG. 4. TDOA vs click time scatterplot for the single sperm whale dataset demonstrating persistent lines corresponding to direct-direct arrival associations, reflected-reflected associations, and mis-associated clicks (direct with surface reflected and vice versa). Also evident toward the end of the recording is a sudden offset in TDOA, corresponding to a clock offset on one of the hydrophones.

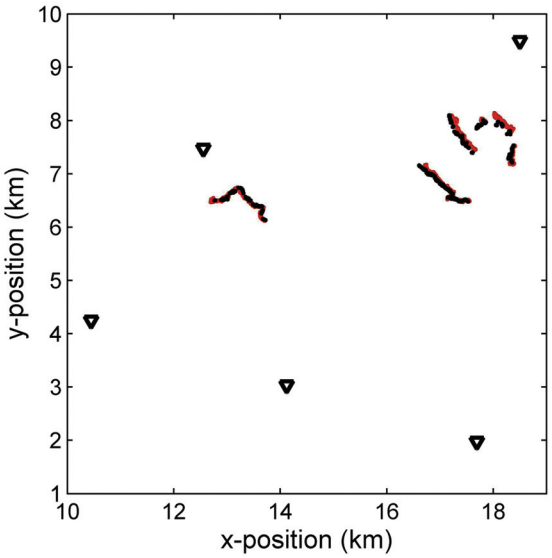


FIG. 5. TDOA (black dots) and TOA (red dots) tracking results. All tracks progress from right to left.

Redistribution subject to ASA license or copyright; see http://acousticalsociety.org/content/terms. Download to IP: 128.171.57.189 On: Wed, 16 Oct 2013 20:25:08

instead of marginals to reduce computational load). This approach, which is outlined in more detail in Nosal and Frazer,[32] can be thought of as measuring the width of the peak in the likelihood surfaces along the x, y, and z directions. For horizontal position estimate, 90% credible interval widths were less than 40 m for clicks heard on five receivers and 70 m for clicks heard on four receivers. 90% credible intervals for depth were less than 500 m for clicks heard on five receivers and less than 800 m for clicks heard on four receivers.

D. TOA

The multiple sperm whale dataset used here proved computationally challenging for the TOA localization method because the number of animals was close to the number of hydrophones (four animals, five hydrophones). As a result the likelihood surface for this dataset had so many spurious peaks that it was computationally prohibitive to find all peaks from which persistent, slowly varying source positions could be tracked. In datasets with fewer animals, more hydrophones, and/or longer ICIs, the localize-then-track approach can be useful. However, for the dataset used here, the two steps of peak/location finding and tracking were combined to make the problem more tractable.

In the combined localization/tracking approach, the search space was divided into cells with a spacing of 20 m in space and 10 ms in time. A systematic search was run over all cells, starting with the first time step and running over all space steps before moving onto the next time step. For each cell that did not yet have an identified track running through it, a constrained Nelder–Mead multidimensional optimizer was used to find the best candidate position within the cell. The best position corresponded to the largest likelihood value using Eq. (7) with the best four phones. The search used the same sound speed profile, look-up table, and interpolation method as the TDOA method. If the likelihood value at the best candidate position exceeded a threshold value of 0.95, the point was flagged as a potential track start point. From this start point, another search was conducted to locate the likelihood peak with value above the 0.95 threshold nearest in time to the start point and within a physically realistic distance to the start point. Using 10 km/h as a maximum swim speed and an inter-call-interval of 5 s (to allow for a missed click detection), the search was constrained to within 5 s and 15 m of the start point. If a second location was identified within the constraints, it was added to the track and the search was repeated starting at the second point. This process was repeated until no further points were found. A track was retained if it contained at least ten points. The resulting tracks are plotted in Fig. 5.

Uncertainty estimates in animal position and click time were obtained in an analogous manner to that used for the TDOA method. 90% credible intervals were less than 15 ms for click time, 40 m for horizontal position estimate, and 300 m for depth.

V. DISCUSSION

Unfortunately there is no way to independently verify the animal tracks obtained here. However, there are a few sanity checks that can be applied to increase confidence in the tracks. First, inter-click intervals established via the TOA tracking method are slowly varying as expected for sperm whale clicks (Fig. 6). Second, the average swim speeds obtained from position estimates were 6.1 km/h using the TDOA method and 5.5 km/h using the TOA method, which are reasonable values for sperm whale swim speeds.[47] Third, the tracks obtained here compare closely with those obtained independently by others working on the same dataset.[13,34] Moreover, during the recording period observers from the Woods Hole Oceanographic Institution saw three whales in the vicinity of the hydrophones,[45] and more may have been present; this agrees with the four animals tracked here.

The tracks obtained using the TDOA are very similar to those obtained using the TOA method. One important difference is that the TOA tracks are more finely sampled (i.e., they contain more position estimates). There are two reasons for this difference: (1) the TDOA method discretizes time, obtaining at most one position per animal per time bin and (2) the implementation of the TDOA method relied on TDOA tracks that had been extracted; a step in which potentially valid TDOA could have been rejected. In contrast, the TOA method produces a position estimate for every click heard on four or more hydrophones which results in more finely sample tracks. For the left-most animal, for example, the fact that nearly all clicks were tracked is evident from the derived ICIs (Fig. 6); if more clicks were missed, ICIs would commonly have values greater than about 1.5 s.

The horizontal position errors associated with the location estimates are likely over-estimates because of the way in which uncertainly was included. Error bars could be reduced by estimating the uncertainly [and resulting standard deviation for use in Eqs. (6) and (7)] on a hydrophone-by-hydrophone and click-by-click basis instead of using a worst-possible scenario taken over all clicks as was done here. The TOA method gave smaller errors in location because some of the uncertainty was taken up by the estimated click times. An even more complete implementation

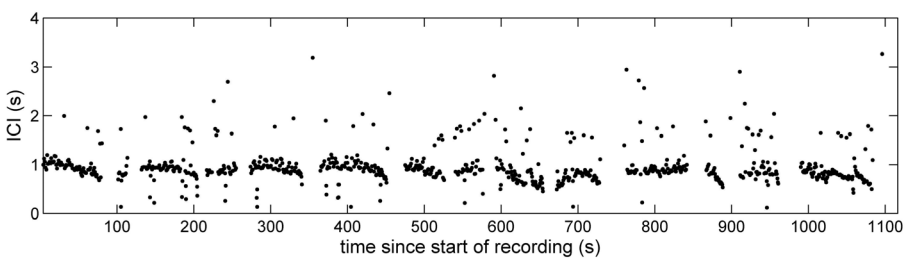


FIG. 6. ICIs obtained from TOA position estimates for the leftmost whale in Fig. 5. ICIs are slowly varying around 1 s, as expected for sperm whale usual clicks.

Redistribution subject to ASA license or copyright; see http://acousticalsociety.org/content/terms. Download to IP: 128.171.57.189 On: Wed, 16 Oct 2013 20:25:08

could also include other uncertainties in the inversion[39–41] (such as SSP, receiver position, and hydrophone timing offset), although doing so greatly increases the computational complexity of the problem by increasing the parameter space. Because all hydrophones were on the seafloor (hence at a similar depth), depth estimates obtained here had much larger errors associated with them than horizontal position estimates did. Using multi-path arrivals could have improved depth estimates by effectively increasing the size of the array.

Run times for both processors were far from real-time; complete processing of the 20 min dataset took about 30 min using the TDOA method over 2 h TOA method on a multicore GPU (Graphical Processing Unit) workstation. The system has a cluster of 448 GPU cores (NVIDIA Tesla) and twelve CPU (Central Processing Unit) cores (Intel Xeon) with 24 GB of SDRAM and a SSD (Solid State Drive) for faster read/write operations. Processing was performed using MATLAB in conjunction with JACKET CUDA software to leverage the GPU cores. The optimization methods used here were chosen for relative ease in implementation (by using the many useful functions available in MATLAB). Very little effort was dedicated to optimizing the implementations or to finding more efficient methods, which is one place that significant improvements could be made in the implementation of the trackers presented here.

Many applications favor lower computational burden and faster run times over high precision and finely sampled tracks. In these cases, it is usually advantageous to pre-associate and prune TDOAs/TOAs when it is possible to do so. By doing this, as many of the spurious peaks as possible from the resulting likelihood surfaces are eliminated which simplifies the optimization (peak finding) process. In fact, the TDOA tracking step used in Sec. IV C is really a TDOA pruning step used to reduce the computational burden associated with the high number of spurious peaks that would be present had all TDOAs in the scatterplot been used. Ultimately, a two- (or multi-)step approach to the tracking problem will probably be the most useful, in which pre-association and pruning are used when possible in a first step and the methods presented in this paper are used in subsequent steps. This approach could possibly incorporate the use of location estimates from earlier steps as *a priori* information on location for use in the subsequent steps (similar to the approach that Baggenstoss[13] took in initializing his tracker with a location estimate from one group of associated TDOAs).

If, on the other hand, very fine track sampling and estimates are desired and computational burden is irrelevant (e.g., when post-processing data in detail for biologically relevant information), it could be useful to include as many TDOAs/TOAs as available. The extreme case would use no pre-association or pruning step whatsoever. In the sperm whale dataset analyzed here, lower detection thresholds would give more false detections (hence more spurious peaks) but also more true detections (hence more location estimates).

As always in practical applications, there is a tradeoff between the accuracy and power of a method on one hand and computational demand and model complexity on the other. The methods presented here lean toward the side of higher power at the expense of higher computational demand, but are flexible enough to allow movement between the two extremes for different applications. The applicability and specifics of any method change on a case-by-case basis, depending (among other things) on the source and environmental characteristics as well as background noise levels and numbers of calling animals. Hopefully the multiple-animal model-based TDOA and TOA methods developed here will prove useful in some cases, such as when associate-then-track methods fail, when highly detailed tracks are desired, or when a combination of methods is the appropriate approach for a problem.

ACKNOWLEDGMENTS

This work was funded by Michael Weise and James Eckman of the Office of Naval Research, Grants Nos. N000140811142 and N000141210206. Data were provided by the Naval Undersea Warfare Center; thanks especially to D. Moretti, R. Morrissey, and N. DiMarzio for help with the dataset and for insightful conversations. Computing resources were provided through a SOEST faculty startup package for EMN.

[1]W. A. Watkins and W. E. Schevill, "Sound source localization by arrival-times on a non-rigid three-dimensional hydrophone array," Deep-Sea Res. **19**, 691–706 (1972).
[2]W. A. Watkins and W. E. Schevill, "Four hydrophone array for acoustic three-dimensional location," Technical Report No. 71, Woods Hole, MA, 1971.
[3]W. C. Cummings and D. V. Holliday, "Passive acoustic location of bowhead whales in a population census off Point Barrow, Alaska," J. Acoust. Soc. Am. **78**, 1163–1169 (1985).
[4]D. H. Cato, "Simple methods of estimating source levels and locations of marine animal sounds," J. Acoust. Soc. Am. **104**, 1667–1678 (1998).
[5]J. Spiesberger, "Passive localization of calling animals and sensing of their acoustic environment using acoustic tomography," Am. Nat. **135**, 107–153 (1990).
[6]K. M. Stafford, C. G. Fox, and D. S. Clark, "Long-range acoustic detection and localization of blue whale calls in the northeast Pacific Ocean," J. Acoust. Soc. Am. **104**, 3616–3625 (1998).
[7]M. A. Mcdonald, J. A. Hildebrand, and S. C. Webb, "Blue and fin whales observed on a seafloor array," J. Acoust. Soc. Am. **98**, 712–721 (1995).
[8]C. W. Clark, W. T. Ellison, and K. Beeman, "Acoustic tracking of migrating bowhead whales," *Proceedings of IEEE Oceans* (1986), pp. 341–346.
[9]R. Payne and S. McVay, "Songs of humpback whales," Science **173**, 585–597 (1971).
[10]H. E. Winn and L. K. Winn, "The song of the humpback whale, *Megaptera novaeangliae*, in the West Indies," Mar. Biol. **47**, 97–114 (1978).
[11]S. M. Van Parijs, C. Lydersen, and K. M. Kovacs, "Vocalizations and movements suggest alternative mating tactics in male bearded seals," Anim. Behav. **65**, 273–283 (2003).
[12]A. Thode, "Tracking sperm whale (*Physeter macrocephalus*) dive profiles using a towed passive acoustic array," J. Acoust. Soc. Am. **116**, 245–253 (2004).
[13]P. M. Baggenstoss, "Separation of sperm whale click-trains for multipath rejection," J. Acoust. Soc. Am. **129**, 3598–3609 (2011).
[14]M. A. Roch, T. S. Brandes, P. Bhavesh, Y. Barkley, S. Baumann-Pickering, and M. S. Soldevilla, "Automated extraction of odontocete whistle contours," J. Acoust. Soc. Am. **130**, 2212–2223 (2011).
[15]A. D. Shapiro and C. Wang, "A versatile pitch tracking algorithm: From human speech to killer whale vocalizations," J. Acoust. Soc. Am. **126**, 451–459 (2009).
[16]J. Starkhammar, J. Nilsson, A. Mats, S. A. Kuczaj, M. Almqvist, and H. W. Persson, "Separating overlapping click trains originating from multiple

Redistribution subject to ASA license or copyright; see http://acousticalsociety.org/content/terms. Download to IP: 128.171.57.189 On: Wed, 16 Oct 2013 20:25:08

individuals in echolocation recordings," J. Acoust. Soc. Am. **129**, 458–466 (2011).
[17]A. Thode, T. Norris, and J. Barlow, "Frequency beamforming of dolphin whistles using a sparse three-element towed array," J. Acoust. Soc. Am. **107**, 3581–3584 (2000).
[18]P. M. Baggenstoss, "An algorithm for the localization of multiple interfering sperm whales using multi-sensor time difference of arrival," J. Acoust. Soc. Am. **130**, 102–112 (2011).
[19]C. W. Clark and W. T. Ellison, "Calibration and comparison of the acoustic location methods used during the spring migration of the bowhead whale, *Balaena mysticetus*, off Pt. Barrow, Alaska, 1984–1993," J. Acoust. Soc. Am. **107**, 3509–3517 (2000).
[20]R. A. Altes, "Detection, estimation, and classification with spectrograms," J. Acoust. Soc. Am. **67**, 1232–1246 (1980).
[21]V. Janik and M. Thompson, "A two-dimensional acoustic localization system for marine mammals," Marine Mammal Sci. **16**, 447–452 (2000).
[22]D. A. Seem and N. C. Rowe, "Shape correlation of low-frequency underwater sounds," J. Acoust. Soc. Am. **95**, 2099–2103 (1994).
[23]P. Giraudet and H. Glotin, "Real-time 3D tracking of whales by echo-robust precise TDOA estimates with a widely-spaced hydrophone array," Appl. Acoust. **67**, 1106–1117 (2006).
[24]J. Spiesberger, "Finding the right cross-correlation peak for locating sounds in multipath environments with a fourth-moment function," J. Acoust. Soc. Am. **108**, 1349–1352 (2000).
[25]P. White, T. Leighton, D. Finfer, C. Powles, and O. Baumann, "Localisation of sperm whales using bottom-mounted sensors," Appl. Acoust. **67**, 1074–1090 (2006).
[26]R. Hirotsu, M. Yanagisawa, T. Ura, M. Sakata, H. Sugimatsu, J. Kojima, and R. Bahl, "Localization of sperm whales in a group using clicks received at two separated short baseline arrays," J. Acoust. Soc. Am. **127**, 133–147 (2010).
[27]L. E. Freitag and P. L. Tyack, "Passive acoustic localization of the Atlantic bottlenose dolphin using whistles and echolocation click," J. Acoust. Soc. Am. **93**, 2197–2205 (1993).
[28]S. M. Wiggins, M. A. McDonald, and J. A. Hildebrand, "Beaked whale and dolphin tracking using a multichannel autonomous acoustic recorder," J. Acoust. Soc. Am. **131**, 156–163 (2012).
[29]B. Miller and S. Dawson, "A large-aperture low-cost hydrophone array for tracking whales from small boats," J. Acoust. Soc. Am. **126**, 2248–2256 (2009).
[30]S. M. Van Parijs, C. W. Clark, R. S. Sousa-Lima, S. E. Parks, S. Rankin, D. Risch, and I. C. V. Opzeeland, "Management and research applications of real-time and archival passive acoustic sensors over varying temporal and spatial scales," Mar. Ecol.: Prog. Ser. **395**, 21–36 (2009).
[31]C. O. Tiemann, M. B. Porter, and L. N. Frazer, "Localization of marine mammals near Hawaii using an acoustic propagation model," J. Acoust. Soc. Am. **115**, 2834–2843 (2004).
[32]E.-M. Nosal and L. N. Frazer, "Sperm whale three-dimensional track, swim orientation, beam pattern, and click levels observed on bottom-mounted hydrophones," J. Acoust. Soc. Am. **122**, 1969–1978 (2007).
[33]C. O. Tiemann, A. M. Thode, J. Straley, V. O'Connell, and K. Folkert, "Three-dimensional localization of sperm whales using a single hydrophone," J. Acoust. Soc. Am. **120**, 2355–2365 (2006).
[34]F. Bénard, H. Glotin, and P. Giraudet, "Highly defined whale group tracking by passive acoustic stochastic matched filter," in *Advances in Sound Localization*, edited by P. Strumillo (InTech, New York, 2011), Chap. 28, pp. 527–544
[35]W. Zimmer, *Passive Acoustic Monitoring of Cetaceans* (Cambridge University Press, Cambridge, 2011).
[36]R. Aubauer, M. O. Lammers, and W. W. Au, "One-hydrophone method of estimating distance and depth of phonating dolphins in shallow water," J. Acoust. Soc. Am. **107**, 2744–2749 (2000).
[37]A. Thode, "Three-dimensional passive acoustic tracking of sperm whales (*Physeter macro-cephalus*) in ray-refracting environments," J. Acoust. Soc. Am. **118**, 3575–3584 (2005).
[38]E.-M. Nosal and L. Neil Frazer, "Track of a sperm whale from delays between direct and surface-reflected clicks," Appl. Acoust. **67**, 1187–1201 (2006).
[39]S. E. Dosso and M. J. Wilmut, "Bayesian multiple-source localization in an uncertain ocean environment," J. Acoust. Soc. Am. **129**, 3577–3589 (2011).
[40]B. Rideout, "Underwater acoustic localization and tracking of pacific walruses in the Northeastern Chukchi Sea," Master's thesis, University of Victoria, Victoria, British Columbia, 2011.
[41]A. M. Thode, G. L. D'Spain, and W. A. Kuperman, "Matched-field processing, geoacoustic inversion, and source signature recovery of blue whale vocalizations," J. Acoust. Soc. Am. **107**, 1286–1300 (2000).
[42]J. L. Spiesberger and M. Wahlberg, "Probability density functions for hyperbolic and isodiachronic locations," J. Acoust. Soc. Am. **112**, 3046–3052 (2002).
[43]J. L. Spiesberger, "Hyperbolic location errors due to insufficient numbers of receivers," J. Acoust. Soc. Am. **109**, 3076–3079 (2001).
[44]M. Wahlberg, B. Møhl, and P. Teglberg Madsen, "Estimating source position accuracy of a large-aperture hydrophone array for bioacoustics," J. Acoust. Soc. Am. **109**, 397–406 (2001).
[45]O. Adam, J.-F. Motsch, F. Desharnais, N. DiMarzio, D. Gillespie, and R. C. Gisiner, "Overview of the 2005 workshop on detection and localization of marine mammals using passive acoustics," Appl. Acoust. **67**, 1061–1070 (2006).
[46]R. P. Morrissey, J. Ward, N. Dimarzio, S. Jarvis, and D. J. Moretti, "Passive acoustic detection and localization of sperm whales (*Physeter macrocephalus*) in the tongue of the ocean," Appl. Acoust. **67**, 1091–1105 (2006).
[47]N. Jaquet, S. Dawson, and L. Douglas, "Vocal behavior of male sperm whales: Why do they click?," J. Acoust. Soc. Am. **109**, 2254–2259 (2001).

Redistribution subject to ASA license or copyright; see http://acousticalsociety.org/content/terms. Download to IP: 128.171.57.189 On: Wed, 16 Oct 2013 20:25:08