

**REPORT DOCUMENTATION PAGE***Form Approved*  
**OMB No. 0704-0188**

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Washington Headquarters Service, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington, DC 20503.

**PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

<b>1. REPORT DATE (DD-MM-YYYY)</b> 06-01-2015		<b>2. REPORT TYPE</b> Monitoring report		<b>3. DATES COVERED (From - To)</b> 1 Jan 2013 - 31 Dec 2014	
<b>4. TITLE AND SUBTITLE</b> Automated acoustic localization and call association for vocalizing humpbackwhales on the Navy's Pacific Missile Range Facility				<b>5a. CONTRACT NUMBER</b>	
				<b>5b. GRANT NUMBER</b>	
				<b>5c. PROGRAM ELEMENT NUMBER</b>	
				<b>5d. PROJECT NUMBER</b>	
<b>6. AUTHOR(S)</b> Tyler A. Helble Glenn R. Ierley Gerald L. D'Spain Stephen W. Martin				<b>5e. TASK NUMBER</b>	
				<b>5f. WORK UNIT NUMBER</b>	
				<b>8. PERFORMING ORGANIZATION REPORT NUMBER</b>	
<b>7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)</b> SPAWAR Systems Center Pacific 53366 Front Street, San Diego, CA Scripps Institution of Oceanography 8635 Discovery Way La Jolla, CA				<b>10. SPONSOR/MONITOR'S ACRONYM(S)</b>	
<b>9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)</b> Commander, U.S.Pacific Fleet 250 Makalapa Dr. Pearl Harbor, HI				<b>11. SPONSORING/MONITORING AGENCY REPORT NUMBER</b>	
<b>12. DISTRIBUTION AVAILABILITY STATEMENT</b> Approved for public release; distribution is unlimited					
<b>13. SUPPLEMENTARY NOTES</b>					
<b>14. ABSTRACT</b> Time difference of arrival (TDOA) methods for acoustically localizing multiple marine mammals have been applied to recorded data from the Navy's Pacific Missile Range Facility in order to localize and track humpback whales. Modifications to established methods were necessary in order to simultaneously track multiple animals on the range faster than real-time and in a fully automated way, while minimizing the number of incorrect localizations. The resulting algorithms were run with no human intervention at computational speeds faster than the data recording speed on over forty days of acoustic recordings from the range, spanning multiple years. Spatial localizations based on correlating sequences of units originating from within the range produce estimates having a standard deviation typically 10 m or less (due primarily to TDOA measurement errors), and a bias of 20 m or less (due primarily to sound speed mismatch). An automated method for associating units to individual whales is presented, enabling automated humpback song analyses to be performed. (Tursiops truncatus, n = 6), false killer whales (Pseudorca crassidens, n = 3) and short-finned pilot whales (Globicephala macrorhynchus, n = 6). Satellite tags were deployed on five different occasions during this period, with four of the five efforts timed to coincide with SCCs (February 2011, August 2011, February 2012, February 2013). The remaining field effort occurred prior to the July 2012 Rim of the Pacific exercise. Initial analysis of tag and PMRF data revealed temporal and general spatial overlap for eight individuals of three species: bottlenose dolphin, short-finned pilot whale, and rough-toothed dolphin. This initial exposure analysis was restricted to one bottlenose dolphin, one short-finned pilot whale, and two rough-toothed dolphins. Based on photo-identification and association analyses, all tagged individuals are known to be from populations generally resident to the islands of Kaua'i and Ni'ihau. Satellite-tagged animals were exposed to estimated received levels of: 130 to 144 decibels for two rough-toothed dolphins, referenced to					

a pressure of 1 micropascal (dB re: 1 $\mu$ Pa) root mean square, hereafter dB; 149 to 168 dB for a bottlenose dolphin; and 141 to 162 dB for a short-finned pilot whale. The bottlenose dolphin showed no large-scale movements out of the area during sonar exposures despite these relatively high predicted received levels, and the short-finned pilot whale actually moved towards areas of higher exposures during the third day of a 3-day period of regular MFAS use. There are a number of acknowledged limitations in terms of the modeling assumptions and the level of resolution on individual response relative to specific sonar transmissions. However, these results demonstrate that this novel integrated approach of using location data from satellite-tagged individuals and modeling to estimate received levels from acoustic recordings from the PMRF hydrophones is a viable and promising approach to examine both estimated exposure levels and potential large-scale movement reactions of tagged individuals.

**15. SUBJECT TERMS**

Monitoring, marine mammal, adaptive management review, Hawaii Range Complex

**16. SECURITY CLASSIFICATION OF:**

**a. REPORT**  
Unclassified

**b. ABSTRACT**  
Unclassified

**c. THIS PAGE**  
Unclassified

**17. LIMITATION OF ABSTRACT**  
UU

**18. NUMBER OF PAGES**  
10

**19a. NAME OF RESPONSIBLE PERSON**  
Department of the Navy

**19b. TELEPHONE NUMBER (Include area code)**  
808-471-6391

# Automated acoustic localization and call association for vocalizing humpback whales on the Navy's Pacific Missile Range Facility

Tyler A. Helble<sup>a)</sup>

SPAWAR Systems Center Pacific, 53560 Hull Street, San Diego, California 92152-5001

Glenn R. Ierley and Gerald L. D'Spain

Scripps Institution of Oceanography, University of California at San Diego, La Jolla, California 92093-0701

Stephen W. Martin

SPAWAR Systems Center Pacific, 53560 Hull Street, San Diego, California 92152-5001

(Received 31 July 2014; revised 25 October 2014; accepted 11 November 2014)

Time difference of arrival (TDOA) methods for acoustically localizing multiple marine mammals have been applied to recorded data from the Navy's Pacific Missile Range Facility in order to localize and track humpback whales. Modifications to established methods were necessary in order to simultaneously track multiple animals on the range faster than real-time and in a fully automated way, while minimizing the number of incorrect localizations. The resulting algorithms were run with no human intervention at computational speeds faster than the data recording speed on over forty days of acoustic recordings from the range, spanning multiple years. Spatial localizations based on correlating sequences of units originating from within the range produce estimates having a standard deviation typically 10 m or less (due primarily to TDOA measurement errors), and a bias of 20 m or less (due primarily to sound speed mismatch). An automated method for associating units to individual whales is presented, enabling automated humpback song analyses to be performed. [<http://dx.doi.org/10.1121/1.4904505>]

[WWA]

Pages: 11–21

## I. INTRODUCTION

Automated localization of marine mammals on the Navy's Pacific Missile Range Facility (PMRF) is important for animal density estimation and behavior studies. Because of the vast amount of stored acoustic data, these automated methods must run faster than real-time in application. Methods for localizing marine mammals using the time of arrival (TOA) or time difference of arrival (TDOA) of incoming transient signals produced by the marine mammals are well established.<sup>1–7</sup> Various implementations of these methods with species-specific considerations are used for localizing certain species of whales on U.S. Navy instrumented training ranges.<sup>8–12</sup> Humpback whales have been problematic to localize using the traditional TOA method, which requires that vocalizations from a given animal be uniquely matched across hydrophones in the array. Humpback songs consist of a sequence of discrete sound elements, called units, that are separated by silence.<sup>13</sup> Units are typically detected from humpbacks every few seconds on the range hydrophones. Arrival times for a given unit at the hydrophones may differ by up to 10 s across the array. Units from an individual are often repeated in a phrase, moreover different individuals may make similar units. Unique association of units across hydrophones is thus challenging. A TDOA method is hence more appropriate and here

implemented by correlating *sequences* of units between pairs of hydrophones. This method is facilitated with use of the generalized power-law (GPL) detector<sup>14</sup> and enhanced with a spectral “templating” procedure to characterize individual units by extracting a fundamental for each unit and setting the remainder of the unit spectrogram to zero. Cross-correlations of sequences of these unit templates allow localization of multiple animals concurrently with an incorrect localization rate of 2% or less. The techniques used are broadly similar to those described in the multiple animal TDOA method in Sec. III A by Nosal.<sup>7</sup> However modifications were made to eliminate both the need for post-processing (thus allowing for real-time localization) and the assumption that animals vocalize frequently enough to produce traceable track lines (a track is defined as a sequence of localizations that can be attributed to one animal). These modifications assume that the number of vocalizing animals in the monitored areas is moderate to low: the algorithms can localize three marine mammals simultaneously in any subarray, with the ability to localize additional marine mammals at the expense of a reduced number of localizations per individual. The algorithms described are capable of operating in real-time on 14 hydrophones. When processing recorded data the algorithms operate on the 14 hydrophones at a rate approximately five times faster than real-time.

In addition to describing the localization methods suitable for real-time processing, a post-processing technique is also described in which information from the localization process is used to assign each unit in the spectrogram to

<sup>a)</sup>Author to whom correspondence should be addressed. Electronic mail: [tyler.helble@navy.mil](mailto:tyler.helble@navy.mil)

individual whales. This added processing step is necessary for call association because the sequences of units originally used for localization may contain units from several whales. Call association can be particularly advantageous to those interested in the biological significance of song and social sound vocalizations in relation to conspecific interactions. Manual annotation of humpback song and social sounds is a laborious and difficult process and therefore automating the majority of the process is beneficial.

The objective of this paper is to describe a robust TDOA localization technique and related call association process focusing on humpback whale vocalizations. While not discussed, the methods are generally applicable to other vocalizing whale species for wide baseline array configurations if the incoming signals can be concurrently detected on four or more hydrophones. Section II A describes the methods used for vocalization detection and feature extraction, Sec. II B describes the cross-correlation techniques used to calculate the TDOAs, and Sec. II C describes the model-based approach used to convert TDOAs into position fixes. Section II D discusses the uncertainties and limitations of the localization process, and includes an optimization analysis for selecting parameters used to minimize localization errors. Section III A details the localization results for two humpback whales transiting through the PMRF range, and the associated uncertainties in the position fixes. Section III B describes the call association process for the same two whales. The final section summarizes the conclusions from this work.

## II. METHODS

The PMRF range is located off the west coast of the island of Kauai in the Hawaiian Islands. Thirty-one time-synchronized hydrophones from the PMRF underwater range have been recorded on a sample basis of approximately two days a month over the past several years, with additional days of recordings associated with U.S. Navy mid-frequency sonar training events. Hydrophone data was initially sampled at 96 kHz and later down sampled to 10 kHz. Of these 31 hydrophones, 14 offshore hydrophones were selected for localization purposes, ranging in depth from 3150 to 4700 m, and covering a rectangular-shaped grid approximately 11 km to the east/west and 52 km to the north/south. The 14 hydrophones were subdivided into four subarrays (A, B, C, D), each containing five hydrophones as shown in Fig. 1. The TDOAs are computed between the center hydrophone of each subarray and the nearest four corner hydrophones. The maximum allowable time delay between the center hydrophone and each adjacent hydrophone in the subarray is limited to the direct path propagation time between them. The subarray configuration was chosen such that a direct path solution on four hydrophone pairs always exists across the monitored area for the noise conditions present on the PMRF range. Additional hydrophones were not included to reduce computational burden. The process for obtaining whale locations can be subdivided into three steps: detection and feature extraction, cross-correlation of those features to obtain TDOAs, and TDOA-based localization.

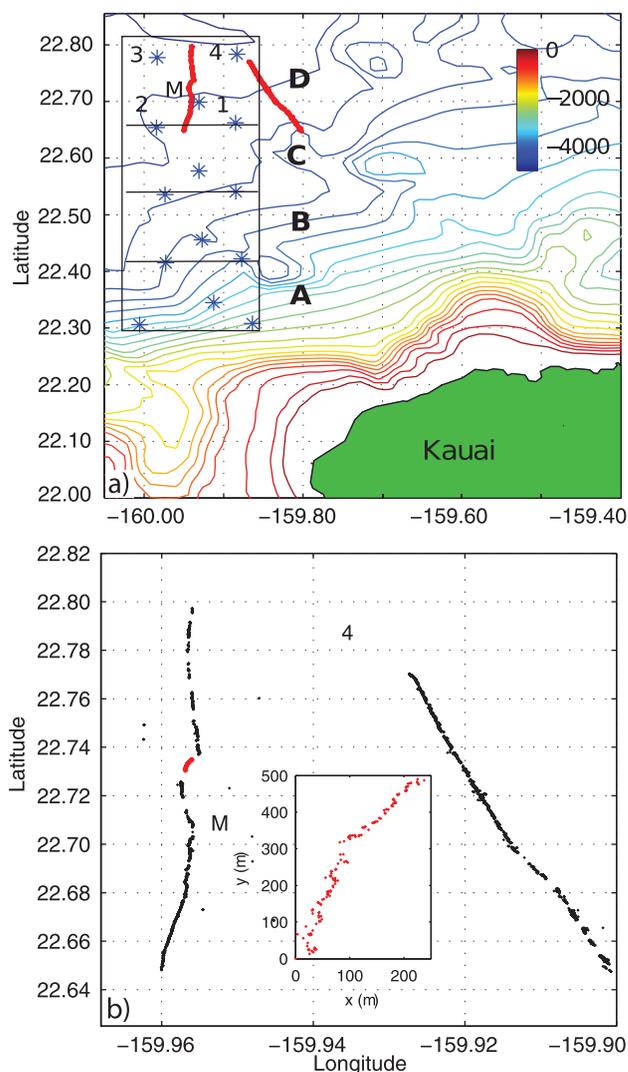


FIG. 1. Approximate positions of PMRF hydrophones illustrating subarrays A–D. The center hydrophone is marked on subarray D (M) and the four adjacent hydrophones (marked 1–4). Position fixes are shown for two humpback whales transiting through subarray D on March 11, 2013 (a), also shown in expanded form (b). The inset shows a detailed portion of the western track (highlighted in red), revealing tightly clustered localization fixes.

### A. Detection and feature extraction

Detection of humpback song units is accomplished using the generalized power-law detector (GPL).<sup>14</sup> The GPL detector is based on the summation of band-limited spectral content. Unlike the energy detector, the GPL algorithm uses a higher power of the Fourier amplitude, which is appropriate when—as for humpback whales—the signal occupies a limited, but unknown, subset of the total search frequencies over which a signal may occur. The GPL detector outperforms energy detectors for humpback song units and has proven effective in accurately determining the start and end times of humpback units in acoustic records under widely varying ocean noise conditions and signal-to-noise (SNR) ratios.<sup>14–16</sup>

In the detection stage, a 60 s spectrogram is band-limited to the 150 to 1000 Hz frequency range and whitened based on an empirical estimate of the noise level at each frequency,  $\mu_k$ , as defined in Eqs. (10), (11) in Helble *et al.*<sup>14</sup>

The Fourier amplitude at each frequency is then normalized by  $\mu_k$ ; a process equivalent to removing the noise that is time-stationary over the duration of the spectrogram.

Once a unit is detected, a templating procedure is used to determine the relevant spectral features of the unit. The full spectrogram is 60 s in length while the segment identified as a unit is only 1 to 2 s. The whitened Fourier amplitude of that short segment is reshaped as a single column vector and its noise level,  $\mu_0$ , now across *both* frequency and time, is determined using the same algorithm [Eqs. (10), (11) in Helble *et al.*<sup>14</sup>]. All elements of the single column vector are then normalized by  $\mu_0$ .

Elements exceeding  $5\mu_0$  above the referenced noise level,  $\mu_0$ , are set to one, and all the remainder are set to zero. The  $5\mu_0$  cutoff value was chosen based on Monte Carlo simulations discussed later in Sec. II D. After restoring the single column vector to its original matrix dimensions, the binary reduction defines a series of “islands.” The main spectral content of the unit is identified as the single island of largest area and remaining islands associated with the unit are discarded. The largest island is used as a mask which is then applied to the original whitened unit spectrogram leaving a single contour, normally the fundamental. The mask can be applied to other powers of the Fourier amplitude as needed in optimization. Single contours are preferred over multiple contours because they prove more robust during the cross-correlation process; the ability to accurately template the harmonics of the unit is variable among the hydrophones due to propagation effects and varying SNR. Figure 2(a) shows the original spectrogram for the center hydrophone on subarray D and the resulting unit templates for the center hydrophone and the four adjacent hydrophones. The unit templates are combined to create a sequence of units used for the cross-correlation process, discussed in the following subsection.

## B. Cross correlation and TDOA

The generalized TDOA method described in Sec. III A by Nosal<sup>7</sup> operates under the assumption that TDOAs have been established between receiver pairs. However, the TDOA on one receiver pair does not need to be associated with the TDOA of another receiver pair and no effort is made to separate false TDOAs (such as incorrect pairings from multipath or incorrect pairing of calls from different animals). In order to vastly reduce ambiguities in localization without requiring a post-processing step, the method discussed here is more restrictive: the center hydrophone of each subarray acts as the “master” and therefore units detected on the center hydrophone must also be detected on each of the four adjacent hydrophones in order to produce a valid localization. Additionally, sequences of humpback vocalization units, rather than single units, are used in the cross-correlation process in order to minimize peaks in the TDOA that arise from incorrect call associations. These more restrictive parameters can be used because the density of the hydrophones is sufficiently high and the water depths sufficiently deep that direct path transmission to each of the five hydrophones is possible in the monitored area. The

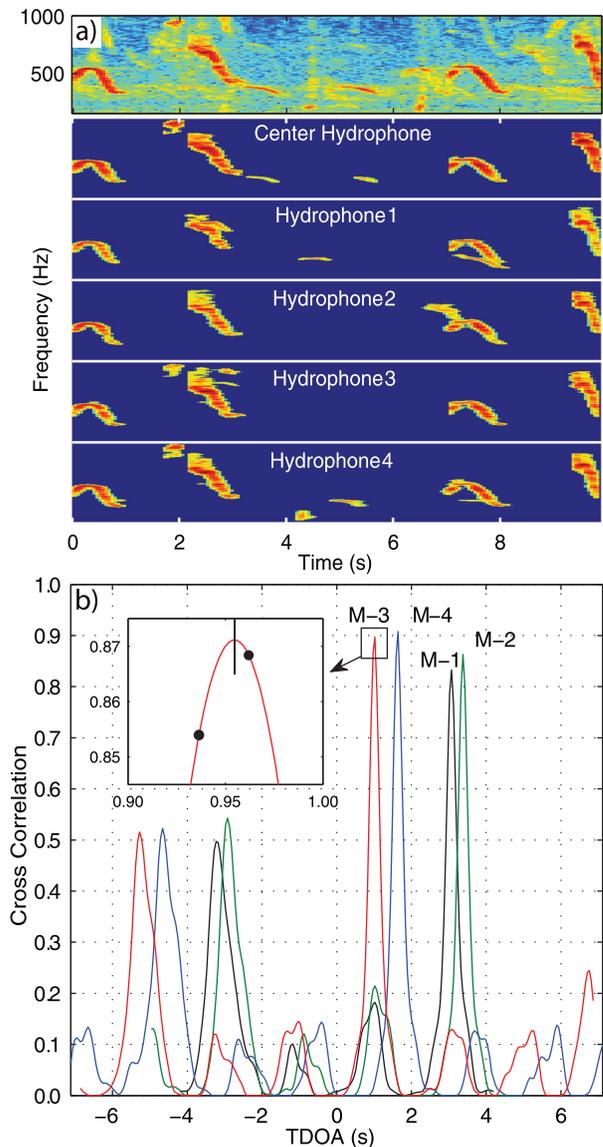


FIG. 2. Original spectrogram (a) for the center hydrophone on subarray D and the resulting template sequences, time aligned according to the highest cross-correlation score between the center hydrophone and each of the four corner hydrophones. Associated cross-correlation (b) revealing peak time delays between the center hydrophone (marked M) and adjacent hydrophones 1–4 for subarray D. The inset shows the interpolation between quantized points in order to obtain a more accurate peak.

sequence of humpback units (rather than single units) can be used in the cross-correlation with minimum degradation in the number of resulting localizations because singing humpbacks produce units every few seconds when they are vocally active, and so whale positions change minimally over the duration of a sequence.<sup>13,17</sup>

Sequences of templated call units as described in Sec. II A are used in the cross-correlation to calculate TDOAs between hydrophone pairs representing varying time windows. An initial sequence is created on the center “master” hydrophone with the desired fixed number of units. The sequence is then cross-correlated with the sequences at each adjacent hydrophone (hydrophones 1–4). The length of the sequence and number of units contained within the sequence at each adjacent hydrophone varies: the sequence contains

all the units detected and templated that occur within the time period of the center hydrophone sequence, plus the direct path travel time padding. After the initial set of computations, the oldest unit is then discarded from the left end of the center hydrophone sequence and a new unit is added on the right (first in, first out), and the next cross-correlation computed. This process is repeated until all units detected on the center hydrophone have been processed. Use of a master sequence containing a fixed number of units, but varying duration, is preferable to reliance on fixed time window sections of spectrograms, but with varying numbers of units, for computing correlations. The latter approach is more prone to false peaks in the TDOA matrices and the estimated time delays between hydrophone pairs are typically less accurate. In principle it is possible to further reduce errors in timing delay estimates by considering all possible pairs of hydrophones within the subarray. However, to remain within the paradigm of TDOA based on sequences of units means identifying the identical sequence on a secondary master hydrophone. Such identification requires unique call identification of individual units and in general that is only feasible once a target animal is already localized.

The accuracy of the average time delay inferred from peak correlation improves with an increase in the sequence size, as documented in Sec. II D. For a moving whale, however, the gain of that statistical improvement must be balanced against the growing divergence between the instantaneous trajectory and the smoothed, time-mean, trajectory predicted by use of the sequence. If these two differ by, say, 25 m, then accuracy in the latter of 5 m is a moot point. Thus, the tradeoff of these two dictates a maximum, preferred, sequence size, which depends upon speed of the whale and the average interval between units. In practice, seven units is a suitable number. For data on the PMRF range, most such sequences last between 10 and 20 s.

The sequences in Fig. 2(a) represents a nine second snapshot in which seven units were detected on the center hydrophone while two humpback whale were traveling through subarray D. Additional vocalizations from distant whales are also present. The sequences are time-aligned according to the highest cross-correlation peak between each hydrophone 1–4 and the center “master” hydrophone. Four of the seven humpback units from the center hydrophone are correctly matched with four units on adjacent hydrophones. The full set of cross-correlations between the center hydrophones and adjacent hydrophones can be seen in Fig. 2(b). The peaks resulting in the correct TDOA stand out prominently compared to the peaks from mismatched units. The second highest set of four correlation peaks are caused by the correlation of a unit with a nearly identical repeated unit produced by the whale approximately seven seconds after the first unit. If single units were used in the cross-correlation, these incorrect peaks would be as prominent as the taller (correct) peaks.

The time series used in the processing was resampled to 10 kHz from the originally recorded 96 kHz sampling rate, and 2048 point fast-Fourier transforms (FFTs) with a Hamming window were used with an overlap of 512 points, resulting in spectral bins with duration of 51.2 ms. The inset in Fig. 2(b) illustrates the discrete values (shown as two

black dots) resulting from the cross-correlation of the spectral templates with temporal bin size of 51.2 ms. Quadratic interpolation about the discrete peaks was used to improve the accuracy of the estimated time delays.

The generalized TDOA method described in Sec. III A by Nosal<sup>7</sup> describes a post-processing step in which the correct TDOAs are selected from the TDOA cross-correlogram by connecting track lines in the image. When multiple animals are present, a feasible real time alternative is to choose the  $N$  most prominent peaks from the TDOA cross-correlogram, allowing up to  $N$  position fixes per sequence. The value of  $N$  cannot exceed the sequence size and in practice should be limited to the number of peaks consistent with position fixes of acceptable accuracy. Note that when animals are calling simultaneously in numbers greater than the chosen value of  $N$ , only  $N$  of the animals will be localized per sequence. Because the sequence on the center “master” hydrophone advances by only one unit at a time, well defined tracks for all of the animals can still be expected.

Figure 3(a) shows the TDOA cross-correlogram between the center hydrophone and hydrophone 4 for the full 3.5 h period in which the two humpback whales traversed the subarray. The highest peaks in the cross-correlogram are a result of the correct cross-correlation, while mismatched correlations are suppressed during the entire period. The corresponding TDOA points to the cross-correlogram are shown in Fig. 3(b) for  $N=3$ , with the highest peaks for each time step shown in red. One then has to test all  $N^4$  combinations of delays but at most  $N$  of these can result in valid localizations. False localizations are rare since the space of valid time delays is a two-dimensional surface and so a random intersection is unlikely. In practice only one or two whales are present within a subarray during the same time period and it suffices to use a single maximum ( $N=1$ ), which confers a notable advantage in computational speedup. However,  $N=3$  is perfectly feasible when needed.

### C. Model-based localization

Localization using the TDOA between hydrophone pairs is accomplished using an established “model-based TDOA”.<sup>7,9,11,18,19</sup> Position fixes are computed using the least-squares difference between the measured and modeled TDOAs, defined as

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

where  $\Delta t_{ij}(k)$  is the  $k$ th measured TDOA that falls within a given time step for receiver pair  $ij$  and  $\Delta \hat{t}_{ij}$  represents the modeled estimate TDOA at position  $\mathbf{w}$ . Applying the “master hydrophone” formulation noted previously,  $i$  is restricted to the center hydrophone of each subarray and  $j$  to the four adjacent hydrophones. Additionally,  $k$  is restricted to the  $N$  largest peaks from the cross-correlation of each sequence. The variance,  $\sigma^2$ , represents errors due to receiver position, measured TDOA, and sound speed profile (SSP). The variances are assumed equal for all hydrophone pairs.

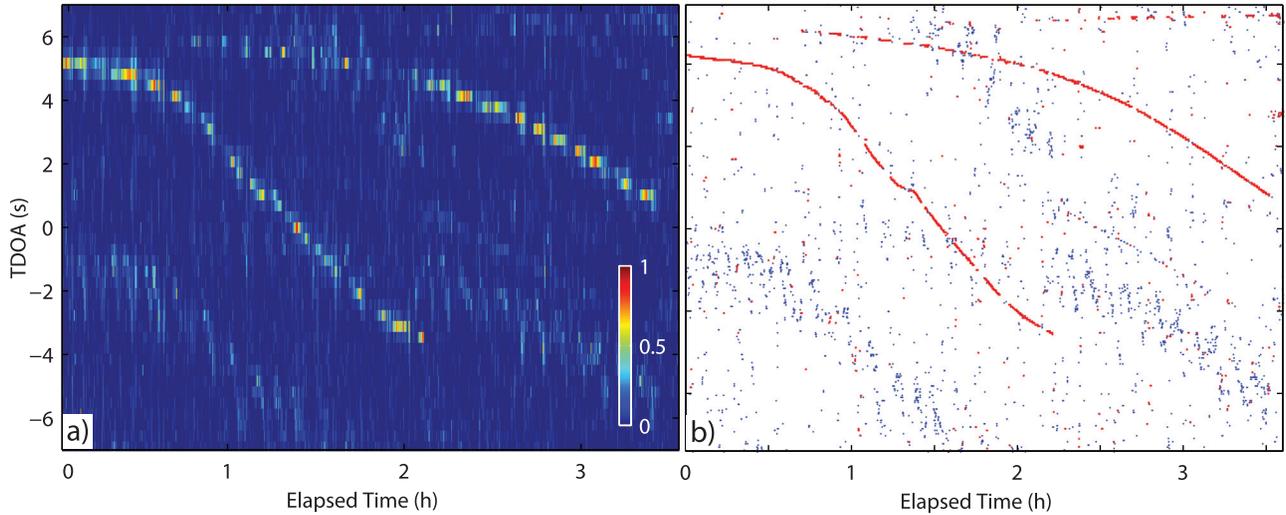


FIG. 3. TDOA cross-correlogram (a) between the center hydrophone and hydrophone 4; the two prominent features represent the TDOA of two whales as they travel through subarray D. The three highest cross-correlation values for each time slice are extracted from the TDOA cross-correlogram and replotted in the right plot (b), representing  $\Delta t_{e4}(k)$  in Eq. (1). The peak values are shown in red, and the second and third highest values shown in blue. The TDOAs represented here are also generated for the center hydrophone to the three other corner hydrophones and the combination of delays are used in Eq. (1) to estimate locations.

The exponential form above is the optimal maximum likelihood estimator on the assumption that the measured delays are independent, identical, and Gaussian distributed.

The quantities  $\Delta \hat{t}_{ij}$  are first computed across the search grid based on estimated direct path travel time of ray paths<sup>20</sup> using a historical SSP. These travel times agreed within 1 ms of travel times predicted by both Bellhop<sup>21</sup> and a range-dependent acoustic model.<sup>15</sup> Distances are estimated using the World Geodetic System (version WGS84) reference ellipsoid.<sup>22</sup> An initial localization is then computed by maximizing  $LS(\mathbf{w})$  in Eq. (1) over a grid with spacing of 50 m. This initial localization is then refined by using the Nelder–Meade optimization,<sup>23,24</sup> allowing the position to vary continuously. A more sophisticated propagation model that incorporates multipath travel times is not needed as the direct path is received at all hydrophones when a whale is calling within the predefined search grid. Eliminating solutions based on multipath arrival reduces localization ambiguities without degradation in localization performance.

As remarked in Sec. II B,  $N=3$  results in 81 candidate sets of TDOAs. Each set of delays with the minimum least-squared value [maximizing Eq. (1)] that meets a threshold criteria is deemed a valid localization. Multiple approaches exist for eliminating the few incorrect localizations that result (for any  $N$ ). The most straightforward approach is to implement a minimum cross-correlation score for each sequence, which ensures that multiple units within a sequence align. A cross-correlation cutoff of 0.4 eliminates all incorrect localizations but at the expense of reducing the number of valid localizations by 20% to 30%. If animals vocalize often enough to ensure that they create traceable track lines and real-time results are not needed, then Nosal’s method<sup>7</sup> can also be used to eliminate the spurious points.

#### D. Sources of uncertainty and limitations

Monte Carlo simulations were used to characterize the performance of the TDOA algorithm as reflected in the

accuracy of timing delays and the resulting position fixes. As localization is a parameter estimation problem, the appropriate metrics of performance are the bias and the variance (or standard deviation) of the estimates. From these simulations emerge optimal values for control parameters in the templating and cross-correlation process that maximize the accuracy of the timing delays and hence that of position fixes as well.

For this purpose, a simulated song was constructed from real recorded humpback units on the PMRF range with a repeated two unit phrase, the first a grunt at 330 Hz lasting 0.75 s, the second a tonal at 530 Hz of 1.5 s duration. Inter-unit spacing was varied between 2.8 and 3.1 s. The SNR of both units was determined by adding white noise of a specified level. The case of “medium-level” noise is defined by the band-limited (restricted to the 150–1000 Hz frequency range over which the GPL detector operates) root-mean-square SNR values of  $-10.6$  dB for the grunt and  $-7.8$  dB for the tonal. The SNR value for the grunt invariably generates a test statistic above threshold for the GPL algorithm. However, about 5% of the time, the detected duration drops below 0.35 s and such units are discarded as false positives. At the SNR level of the tonal, the missed detection rate is about 10%. This higher rate arises both from occasional failure of the test statistic to rise above threshold and/or dropping below the duration limit, when the unit is fragmented in the spectrogram. Figure 4 exhibits instances of all these shortcomings. In practice, real humpback signals originating from within the range always contain SNR values of this level or higher over all noise levels recorded on the range. The templating threshold value of  $5\mu_0$  described in Sec. II A is thus set at an appropriate level for templating nearly all direct-path arrival units originating from within the range.

The major shortcoming of these Monte Carlo runs is that transmission loss on the range is not modeled. While all five hydrophones thus receive identical signals, the noise realizations are independent and hence statistical variability occurs in the detection and templating process between hydrophones.

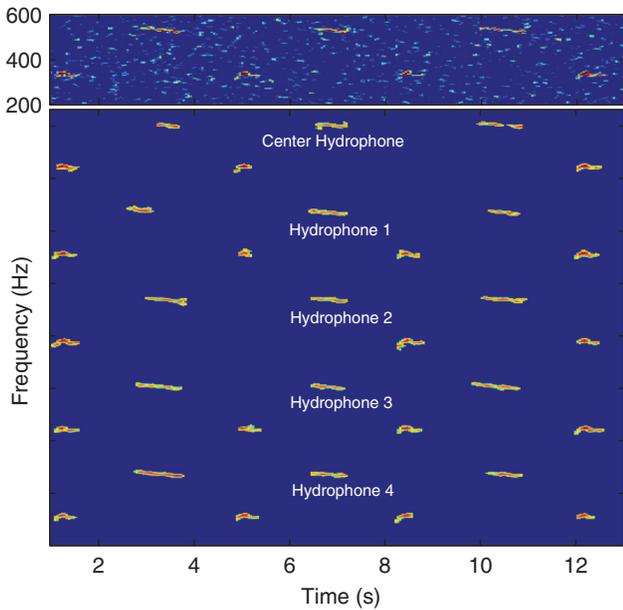


FIG. 4. Example spectrogram and related time-aligned sequences used in the Monte Carlo simulations for the medium-level noise benchmark case (B1). The sequences are aligned according to highest cross-correlation score for the center hydrophone to each of the four corner hydrophones. Band limited spectrograms around the strongest harmonic vary between  $\pm 2$  dB.

For each trial, the location of the vocalizing whale was determined from a two-dimensional Gaussian spatial probability density function (pdf) with  $\sigma_x = \sigma_y = 1$  km and centered at a point 9.25 km NW from the center hydrophone of subarray D. Time delays were determined from the detection, templating, and resulting cross-correlation of sequences of the simulated song units. These empirical TDOAs were substituted into Eq. (1) and the least squares position estimate determined from the best match with predicted direct ray path travel times. The resulting position estimate was then compared against the known vocalization origin point.

Many parameters in the detection, templating, and cross-correlation process affect the accuracy of the timing delays. Based on numerous Monte Carlo simulations, as well as application to real data from the PMRF range, it was determined that the best results are obtained by (1) characterizing units by a single harmonic, (2) basing correlation on (normalized) Fourier amplitude (not amplitude squared) within that harmonic, (3) using quadratic interpolation of the digitized correlation data to refine the peak, (4) using a sequence of seven calls, and (5) using an FFT length of 2048 with an overlap of 75% (bin spacing of 51.2 ms). These parameters define the “benchmark case.” Altering any of conditions (1)–(3), or relaxing (4) or (5) (i.e., reducing the sequence size or decreasing the FFT overlap), all degraded the performance.

While it might be thought that *increasing* the overlap beyond the stated 75% would monotonically improve results, finer temporal resolution at fixed FFT length does not improve the accuracy in determining peak correlation time. Statistics for an overlap of 93.75%—a bin spacing of 12.8 ms—are worse. However, doubling the FFT length to 4096 while increasing the overlap to 93.75%—a bin spacing of 25.6 ms—does improve accuracy, but only slightly, while the CPU time increases substantially.

Table I characterizes performance of the TDOA algorithm for various choices of model parameters in terms of  $\sigma_t$ , the standard deviation for time delay errors,  $\sigma_x$ , the standard deviation of the local Cartesian expansion of latitude, and  $\sigma_y$ , the standard deviation of the local Cartesian expansion of longitude. Further comments on characterizing the error appear shortly. Each case represents 1000 trials.

The benchmark case is shown for medium-level noise (B1), low-level noise (defined to be 3 dB down from medium-level noise) (B2), and for the zero noise limit (B3). The zero noise limit case represents the irreducible, intrinsic, errors associated with the benchmark parameter set (1)–(5). The increase in  $\sigma_t$  above this scales linearly with energy in the noise (a factor of 4 from the low-level to medium-level noise case).

Additionally, errors are shown when the benchmark case Fourier amplitude templates are replaced with Fourier amplitude-squared (energy) (T1), and when the templates are solely based on shape (T2). Note that using only the shape as used by Tiemann *et al.*<sup>11</sup> is notably worse. While results in Table I for (T1) are equivalent to (B1), in application to real data, (T1) yields appreciably fewer position fixes and this observation, rather than the Monte Carlo simulation, is the basis for defining amplitude as the benchmark.

Finally, errors are tabulated for varying sequence size with two calls per sequence, containing both the grunt and tonal (S1), and single call sequences for the grunt (S2) and tonal (S3). Note for the single call, the delay errors are considerably worse for the tonal, which exhibits nonnormal statistics with a fat tail, probably arising from call fragmentation.

For the eight cases reported here, a total of 32 time differences of arrival exist for the four hydrophone pairings. The corresponding sample means all lie within  $\pm 1.5$  times the standard error of the mean, consistent with a uniform assumption of zero bias for the TDOA algorithm, as well as the position fixes subsequently estimated.

The cumulative distribution function (cdf) of benchmark timing delay errors (B1–B3 in Table I) can be fit by the normal form,  $\Phi(t) = 1/2(1 + \text{erf}(t/\sigma_t\sqrt{2}))$ . For the medium-level noise benchmark case  $\sigma_t = 4.85$  ms. This zero bias fit satisfies the two-sided Kolmogorov–Smirnov (K–S) null hypothesis test at a significance level of 0.05 and hence justifies use of the Gaussian distribution in the expression for the maximum likelihood estimator in Eq. (1). Zero bias in the

TABLE I. The standard deviation for time delay errors  $\sigma_t$  and the standard deviation of the local Cartesian expansion of longitude  $\sigma_x$  and latitude  $\sigma_y$  for various cases used in the Monte Carlo simulations.

Case	Description	Noise level	$\sigma_t$	$\sigma_x$	$\sigma_y$
B1	Benchmark	Medium	4.85 ms	6.43 m	5.58 m
B2	Benchmark	Low	1.69 ms	2.25 m	1.96 m
B3	Benchmark	None	0.61 ms	0.98 m	0.64 m
T1	Energy	Medium	4.84 ms	6.52 m	5.68 m
T2	Shape	Medium	9.53 ms	12.95 m	11.16 m
S1	Two unit	Medium	11.40 ms	15.84 m	13.12 m
S2	Single grunt	Medium	9.06 ms	11.48 m	10.76 m
S3	Single tonal	Medium	42.61 ms	58.44 m	48.75 m

timing errors implies zero bias in estimates of longitude and latitude from Eq. (1). The zero bias spatial expectation is confirmed by the K–S test for the respective normal form cdf benchmark fits with  $\Phi(x)$  using  $\sigma_x = 6.43$  m for latitude and  $\Phi(y)$  with  $\sigma_y = 5.58$  m for longitude in the medium-level noise case. It is useful to combine the last two standard deviations as a single measure,  $\sigma_d = \sqrt{(\sigma_x^2 + \sigma_y^2)}/2 = 6.02$  m. The related pdf for random vector lengths in the plane is the general Rayleigh distribution with variance  $(4 - \pi)\sigma_d^2/2 = 15.58$  m<sup>2</sup> and the mean given by  $\bar{d} = \sigma_d \sqrt{\pi/2} = 7.55$  m. These values are to be compared to a sample variance of 15.46 m<sup>2</sup> and sample mean of 7.54 m. This Rayleigh distribution is plotted in black and green for the medium-level and low-level noise benchmark cases (B1, B2) in Fig. 5, and the purely empirical pdf's for other cases (T1, T2, S1) are shown in red, blue, and purple, respectively. While (T1) and (T2) are also arguably Rayleigh distributions based on the K–S test, case (S1), the two unit sequence, is manifestly not: it has a tail that decays with a controlling factor of  $\exp(-\alpha d)$  rather than  $\exp(-\alpha d^2)$ . The time delay errors in that case also have a long tail, one not modeled by  $\Phi(t)$ .

Without a database of known source locations and accompanying time series, the only way to characterize the distribution of time delay errors is with Monte Carlo simulations. How that distribution translates into errors for latitude and longitude is then directly a function of the array geometry and the sound speed profile. The standard deviations for the errors in position can be calculated by minimizing a quadratic approximation of Eq. (1) based on a first order expansion for  $\Delta \hat{t}_{ij}(\mathbf{w})$ . Therefore, it is possible to obtain localization error estimates over the entire array, rather than just the small region to which the Monte Carlo results were limited. The expected localization errors in latitude and

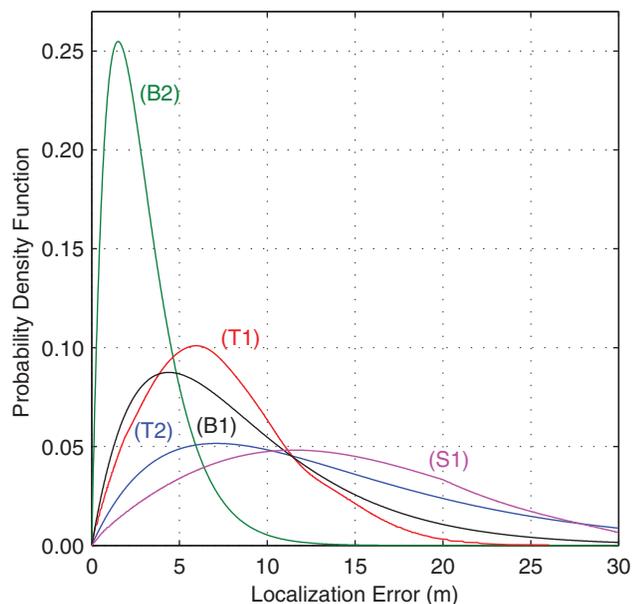


FIG. 5. Probability density function (pdf) localization errors are shown for the benchmark cross-correlation and TDOA parameters for medium-level noise (black) and low-level noise (green) and several alternative cases discussed in Sec. IID. The symbols (B1, B2, T1, T2, S1) correspond to the descriptions in Table I.

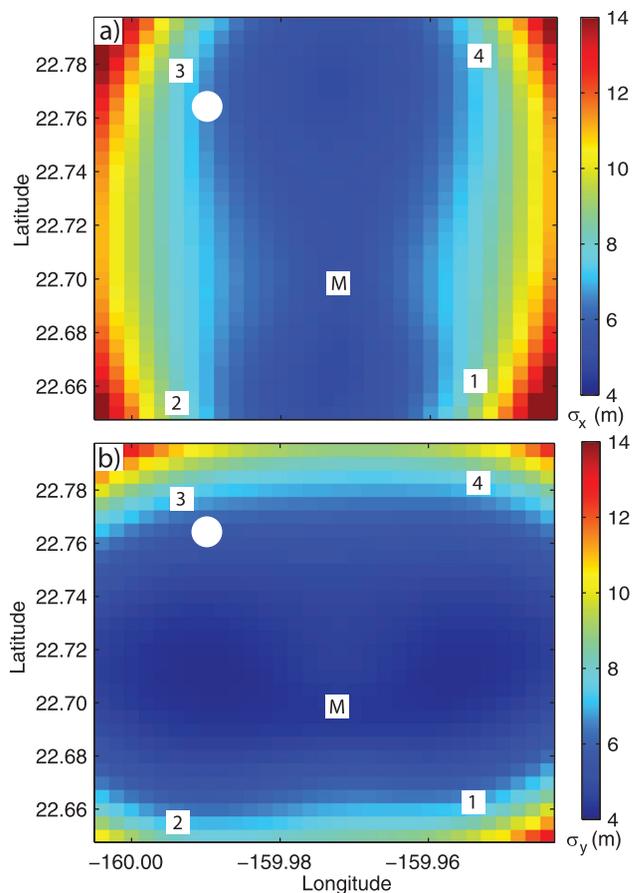


FIG. 6. Theoretical calculation of the standard deviation of horizontal localization errors  $\sigma_x$  (a) and vertical localization errors  $\sigma_y$  (b) as a function of longitude and latitude for correlated random timing delay errors in Eq. (1). Approximate locations for hydrophones (1–4) and center hydrophone (M) are shown for subarray D. The Monte Carlo simulations for distance errors were limited to the 1 km radius patch marked by the white circle.

longitude for subarray D are shown in Fig. 6, with the Gaussian patch for the Monte Carlo simulations marked by a white circle. For the benchmark case (B1) the theoretical predictions using  $\sigma_t = 4.83$  ms are  $\sigma_x = 6.51$  m and  $\sigma_y = 5.67$  m. These agree with the Monte Carlo values of 6.43 and 5.58 m to within expected error for 1000 trials. Since the position errors scale linearly with  $\sigma_t$ , the general patterns in Fig. 6 show the relation of position errors for any of Monte Carlo simulations except S1–S3 relative to their magnitude in the rest of the domain for any noise level. As anticipated, errors are largest in the corners and rise sharply outside the borders of the array, though longitude is more sensitive to the east and west and latitude to the north and south.

All the Monte Carlo results were obtained with a spatially stationary whale. Generally, however, the whale is moving along a fairly linear trajectory at a nearly uniform speed. It can be anticipated that such movement degrades the accuracies noted previously. To address this issue, a simulation of 2000 trials was run for the benchmark case (medium-level noise, seven unit sequence, correlation on amplitude) with the position of the whale at  $t = 0$  chosen as above, but also an azimuth selected from a uniform distribution on  $(0^\circ, 360^\circ)$ . The whale was assumed to travel at

6 km/h along a great circle with the given azimuth. For a sequence of calls that spans 18 s, the change in position is thus 30 m. The standard deviation in timing delay increases from 4.85 to 5.40 ms, the mean position error from 8.49 to 9.00 m. Comparable adjustments can be expected for the other cases, with their rank order unchanged.

Finally, as noted above, the time delay errors are well modeled as identically distributed normal random variables. It turns out that they are not, however, independent, presumably because the center hydrophone is common to all four TDOA estimates. Equation (1) is therefore not the best linear unbiased estimator. Rather one should incorporate the inverse of the covariance matrix in the quadratic form inside the exponential. In general the elements of this matrix would depend upon both position in the subarray owing to path dependence of transmission loss, an effect not modeled in the Monte Carlo simulations, and also background noise level. The issue merits further investigation but one can note that the changes to Fig. 6 based on a simple model covariance matrix for the idealized case of the Monte Carlo simulations are modest, with the qualitative variation across the array unchanged.

The goal of the Monte Carlo simulations was to identify optimal methods for estimating time delays from measured time series. But the other half of the problem, predicting time delays, rests on data with other sources of uncertainty, namely the sound speed profile with depth and range, and hydrophone locations and depths. The sound speed profiles shown in Fig. 7(a) were used to calculate the standard deviation and bias for local Cartesian coordinates  $x$  and  $y$  when a sound speed mismatch is used. The extremal profiles (red and green) represent the variation for 23 measurements taken

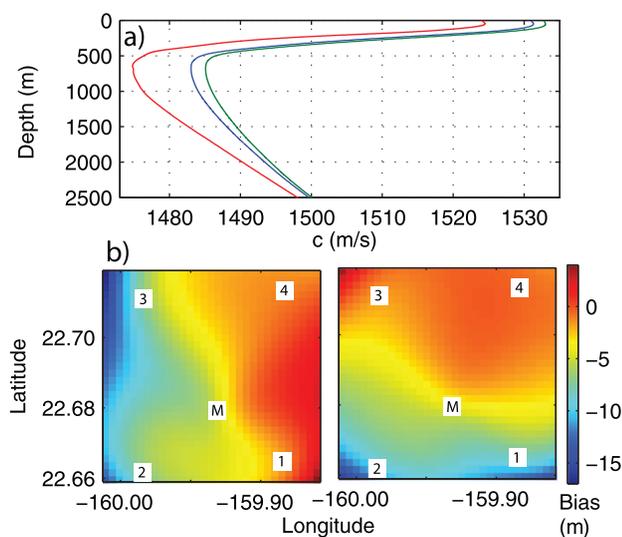


FIG. 7. Representative sound speed profiles (a) showing a February 2013 cast (blue) and maximal variation (red and green) for 28 measurements over multiple years and all seasons at PMRF. The variation between the blue and green curves is representative of typical variation observed over the course of a month on the range. The bias for local Cartesian coordinates  $x$  (left plot) and  $y$  (right plot) as a function of longitude and latitude is shown in the lower plots (b) for subarray D. For the bias shown, the blue sound speed profile is assumed to be the actual *in situ* profile, but the green sound speed profile is used to calculate the predicted time delays. Approximate locations for hydrophones (1–4) and center hydrophone (M) are shown for subarray D.

over all months and spanning several years, while the blue profile is an SSP for February 2013. For the medium-level noise benchmark case (B1), the standard deviations  $\sigma_x$  and  $\sigma_y$  are unchanged when exact measured time delays were computed from the blue SSP and the predicted time delays were based on the green SSP. A second effect of uncertainty in sound speed is the bias in position fixes. This point is illustrated in Fig. 7(b), which shows the bias in the local Cartesian coordinates  $x$  and  $y$  that results from minimizing Eq. (1) with exact measured time delays computed from the blue SSP and predicted time delays based on the green SSP. The displacements in this case are of the same order as the errors  $\sigma_x$  and  $\sigma_y$ , but the former vary gradually, on a scale of a few km, while the latter vary from one position fix to the next. This disparity means that detailed features in whale trajectories are preserved and merely displaced with a nearly rigid translation. Note that variation between the blue and green SSPs is typical over the course of a month. The maximal variation over all measured seasons and years, i.e. measured time delays from the red SSP and the predicted time delays from the blue SSP, increases the scale in Fig. 7(b) by a factor of 4.

A second source in bias arises from the uncertainty in depth of the vocalizing animal. The contours for this bias are broadly similar to those in Fig. 7(b). In the specific case of a whale vocalizing at 100 m depth but assumed to be at 5 m depth there results an induced horizontal root-mean-square bias of 6.1 m averaged over the area enclosed by subarray D.

In the case of the PMRF range, hydrophone locations and depths are well characterized, so errors in hydrophone positions were not explored. In experiments where either sound speed or hydrophone locations are less well known, note that one can bootstrap to good effect by, for example, allowing the locations of  $n - 1$  hydrophones to vary while holding one fixed and maximizing  $LS(\mathbf{w})$  in Eq. (1) for  $2n$  (or more) position fixes simultaneously. When both sound speed and hydrophone locations are poorly known, bootstrapping will generally yield families of solutions for hydrophone locations and sound speed, rather than a single optimum. As in the case of sound speed mismatch the standard deviations  $\sigma_x$  and  $\sigma_y$  are unchanged.

### III. RESULTS

#### A. Localization

Figure 1 shows the track lines of the two humpback whales that are associated with the template and TDOA examples in Figs. 2 and 3. A total of 3500 valid localizations were computed over the course of 3.5 h. Post-processing (discussed in Sec. III B) reveals that 877 localizations can be assigned to individual units for the western track, and 1060 units can be attributed to the eastern track. Because sequences of units are used to calculate the localizations, an individual unit can contribute to multiple localizations, and hence the higher total localization tally. The average estimated inter-call interval for the western track based on the detected units is 3.3 and 3.2 s for the eastern track. Manual inspection of the original spectrograms over the same period reveals an inter-call interval of song units of approximately

3 s for both whales, suggesting that valid localizations are calculated for nearly every unit produced. A few clusters of incorrect localizations can be seen in Fig. 1, totaling 25 points over a period of 22 h. As noted earlier, such incorrect localizations can be eliminated by implementing a cross-correlation minimum of 0.4 during the calculation of TDOAs, but the number of valid localizations drops from 3500 to 2500. Even with this reduction, both track lines remain well defined with multiple localizations every minute during intervals of vocalization. The incorrect localizations could also be easily removed by implementing one or both of the post-processing steps outlined by Nosal,<sup>7</sup> with the advantage of keeping all valid points along the track line.

The methods outlined in this paper were used to process a total of 40 days of recordings on the PMRF range over the months of December through May, spanning the years from 2011 through 2014. Thirty-one unique humpback track lines were found in the recordings. Manual inspection of the TDOA cross-correlogram revealed that all calling humpbacks in the vicinity were localized consistently. A surprisingly large proportion of detected units originate from off-range locations. While exact position fixes cannot be calculated, analysis suggests these calls originate from near-shore and potentially propagate up to 60 km in some cases. If analysis were done on single hydrophones within the range, the animal density could easily be overestimated in the study area, due to the non-random distribution of animals. The processing of all four subarrays was accomplished five times faster than real-time on a standard dual-core computer with 2.2 GHz processors. A slight time delay is required to amass enough units to construct the sequence, but on average, this delay is on the order of 20 s or less for actively calling whales.

While no data exist to ground-truth the localization coordinates produced from the recorded data, some aspects of the results indicate that the localization accuracies are consistent with those estimated from the Monte Carlo simulations. The inset in Fig. 1(b) shows very tightly clustered localization points along the presumed track. Fitting a trajectory through this trajectory, it was found that no point deviated by more than 100 m from the track, and the standard deviation from the track line was 17 m. These results are consistent with the error estimates predicted from the Monte Carlo simulations. Additionally, during the development of the sequencing process, various sequence sizes were considered. The sequence size can be increased beyond seven,

though with diminishing return. A trial computation with a twelve-call sequence for real data from the PMRF range does show perceptibly tighter grouping, particularly in the eastern (off-range) trajectory of Fig. 1(b). However, of some 3500 position fixes based on the same initial unit, latitudes for the twelve-call sequence show an 8.7 m bias to the north and 1.02 m bias to the west. The twelve-call sequences last an average of 10.6 s longer and the northward bias is consistent with a mean northward velocity of 5.9 km/h. The whale on the western track averages 6.7 km/h northward, that on the eastern track averages 5.2 km/h northward. An approximately equal number of calls are detected from each track and hence the overall bias is accounted for to within a few percent by the average these two speeds. The westward bias is similarly explained. Independent localizations for the western track were computed using both subarrays C and D as the whale transited across the subarray border. A total of 357 localizations from subarray C were compared to positions from subarray D interpolated for the same time (referenced to the whale's position). The comparison yielded a localization agreement with standard deviation of  $\sigma = 9.8$  m. The observed cluster tightness, the velocity estimates from sequence comparison, and the agreement of independent localizations all provide excellent evidence that the Monte Carlo simulations with time delay errors of 5–10 ms give a realistic estimate of expected errors. Additionally, the humpback transiting speeds noted above are consistent with observational data for transiting humpback whales.<sup>25</sup>

## B. Call association

Once whale tracks have been established, it is possible to post-process the acoustic data and assign humpback song units within a spectrogram to individual singers. The general procedure is to first calculate the expected TDOA between the center hydrophone to adjacent hydrophones for all locations along the track line. Next, the cross-correlation score for each individual unit is calculated between the master hydrophone and each adjacent hydrophone in the vicinity of the expected delay, allowing for a variation of  $\pm 5$  ms. If the unit has a cross-correlation score of 0.4 or higher on at least two hydrophone pairings, then the unit is assigned to the individual whale on the track of interest. Figure 8(a) shows the original spectrogram on the center hydrophone for subarray D, containing song from the two whales whose tracks are shown in Fig. 1, with the song from at least one more distant

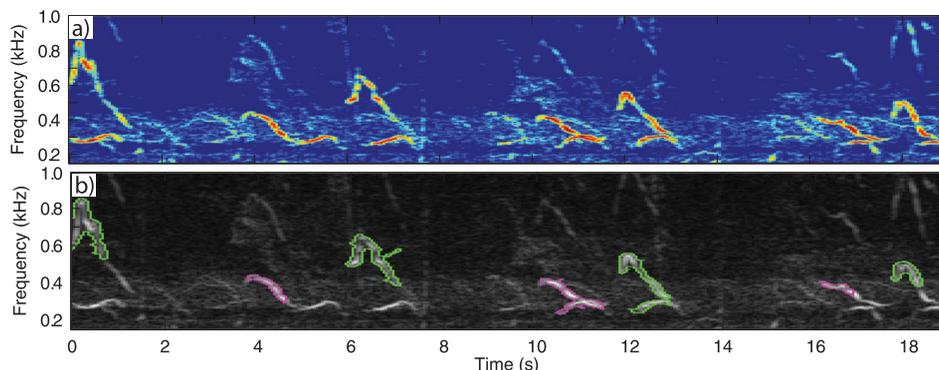


FIG. 8. Spectrogram from the center hydrophone of subarray D (a) recorded during vocalization of two humpback whales as they transit through the subarray (shown in Fig. 1), with additional distant whale vocalizations present. The same spectrogram (b) shown with automated color contours drawn representing the whale from the western track (green) and the whale from the eastern track (purple).

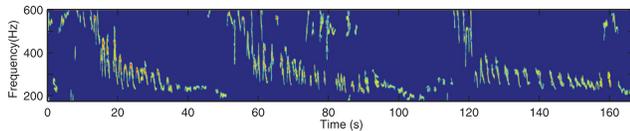


FIG. 9. Automated extraction of humpback units (shown in template form) relating to the western whale track from Fig. 1. Inter-call intervals have been suppressed so that the song sequencing is clearer (actual elapsed time is approximately 10 min).

whale also present. Figure 8(b) shows a grayscale image of the same spectrogram, with the detected unit contours from the templating procedure discussed in Sec. II A highlighted in green for the western track whale and magenta for the eastern track whale. The low frequency units from the third more distant whale occasionally overlap in time and frequency with the units produced by the other two whales, and so sometimes a contour is merged across units originating from two separate whales. While not perfect, this automated assignment of most units to individual whales can be helpful for biologists interested in annotating humpback song, or examining the relationship of song production between conspecifics. Figure 9 shows the templates of the song sequence assigned to the western whale, with the time between units removed. Manual analysis shows that approximately 90% of the units produced by the western whale are templated and assigned correctly, and no units from other whales are included (unless a unit overlaps in space/time with a unit originated from the western whale, in which case some of the contour can be included). Using this technique, automatically extracted relevant song information can be used for analysis. Currently, the center, lowest, and highest frequency of each templated unit is automatically saved. This information may prove useful for automatically harvesting large-scale statistics on humpback calling patterns.

#### IV. CONCLUSIONS

The techniques outlined in this paper prove to be effective for localizing humpback whale vocalizations on 14 hydrophones five times faster than real-time on the PMRF range with a predicted high level of spatial accuracy. The localization process is robust over a wide range of environmental and noise conditions, and has been shown to work on data collected in the months of December to May over multiple years. Although not discussed in detail, the GPL detection and templating procedure is general enough to be readily adapted to other types of marine mammal vocalizations, and so the same process for obtaining TDOAs between hydrophone pairs can be ported to other species. The model-based localization method outlined in this paper is built on many of the same principles described in other peer-reviewed publications, and has proven to work well over a variety of species, array configurations, and bathymetric and environmental conditions. The post processing methods outlined for call association could prove helpful for matching vocalizations to individual whales, even in the presence of multiple calling animals with similar vocal patterns. One obvious extension of the call association process is to

automatically obtain cue rates from existing data sets. Obtaining information on cue rates over a variety of social, spatial, temporal, and environmental conditions is a crucial component for calculating animal densities from passive acoustic data.

#### ACKNOWLEDGMENTS

Work was supported by the Office of Naval Research, Code 322 (MBB), Commander, U.S. Pacific Fleet, Code N01CE, and the Naval Facilities Engineering Command Living Marine Resources Program. Additionally, the authors thank Brian Matsuyama at SPAWAR Systems Center Pacific for his assistance with data collection and analysis.

- <sup>1</sup>W. Watkins and W. Schevill, "Sound source localization by arrival times on non-rigid three-dimensional hydrophone array," *Deep-Sea Res.* **19**, 691–706 (1972).
- <sup>2</sup>W. Watkins and W. Schevill, "Four hydrophone array for acoustic three-dimensional location," Technical Report No. 71, Woods Hole Oceanographic Institution, Woods Hole, MA (1971).
- <sup>3</sup>D. Cato, "Simple methods of estimating source levels and locations of marine animal sounds," *J. Acoust. Soc. Am.* **104**, 1667–1678 (1998).
- <sup>4</sup>P. Giraudet and H. Glotin, "Real-time 3D tracking of whales by echorobust precise TDOA estimates with a widely-spaced hydrophone array," *Appl. Acoust.* **67**, 1106–1117 (2006).
- <sup>5</sup>A. Širović, J. Hildebrand, and S. Wiggins, "Blue and fin whale call source levels and propagation range in the Southern Ocean," *J. Acoust. Soc. Am.* **122**, 1208–1215 (2007).
- <sup>6</sup>P. 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).
- <sup>7</sup>E.-M. Nosal, "Methods for tracking multiple marine mammals with wide-baseline passive acoustic arrays," *J. Acoust. Soc. Am.* **134**, 2383–2392 (2013).
- <sup>8</sup>C. Tiemann, S. Martin, and J. Mobley, "Aerial and acoustic marine mammal detection and localization on navy ranges," *IEEE J. Oceanic Eng.* **31**, 107–119 (2006).
- <sup>9</sup>R. Morrissey, J. Ward, N. DiMarzio, S. Jarvis, and D. Moretti, "Passive acoustic detection and localization of sperm whales (*Physeter macrocephalus*) in the tongue of the ocean," *Appl. Acoust.* **67**, 1091–1105 (2006).
- <sup>10</sup>T. Marques, L. Thomas, J. Ward, N. DiMarzio, and P. Tyack, "Estimating cetacean population density using fixed passive acoustic sensors: An example with Blainville's beaked whales," *J. Acoust. Soc. Am.* **125**, 1982–1994 (2009).
- <sup>11</sup>C. Tiemann, M. Porter, and L. N. Frazer, "Localization of marine mammals near Hawaii using an acoustic propagation model," *J. Acoust. Soc. Am.* **115**, 2834–2843 (2004).
- <sup>12</sup>S. Martin, T. Marques, L. Thomas, P. Morrissey, S. Jarvis, N. DiMarzio, D. Moretti, and D. Mellinger, "Estimating minke whale (*Balaenoptera acutorostrata*) being sound density using passive acoustic sensors," *Mar. Mammal Sci.* **29**, 142–158 (2013).
- <sup>13</sup>R. Payne and S. McVay, "Songs of humpback whales," *Science* **173**, 585–597 (1971).
- <sup>14</sup>T. Helble, G. Ierley, G. D'Spain, M. Roch, and J. Hildebrand, "A generalized power-law detection algorithm for humpback whale vocalizations," *J. Acoust. Soc. Am.* **131**, 2682–2699 (2012).
- <sup>15</sup>T. Helble, G. D'Spain, J. Hildebrand, G. Campbell, R. Campbell, and K. Heaney, "Site specific probability of passive acoustic detection of humpback whale calls from single fixed hydrophones," *J. Acoust. Soc. Am.* **134**, 2556–2570 (2013).
- <sup>16</sup>T. Helble, G. D'Spain, G. Campbell, and J. Hildebrand, "Calibrating passive acoustic monitoring: Correcting humpback whale call detections for site-specific and time-dependent environmental characteristics," *J. Acoust. Soc. Am.* **134**, EL400–EL406 (2013).
- <sup>17</sup>H. Winn, T. Thompson, W. Cummings, J. Hain, J. Hudnall, H. Hays, and W. Steiner, "Song of the humpback whale-population comparisons," *Behav. Ecol. Sociobiol.* **8**, 41–46 (1981).

- <sup>18</sup>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).
- <sup>19</sup>J. Spiesberger and M. Wahlberg, "Probability density functions for hyperbolic and isodiachronic locations," *J. Acoust. Soc. Am.* **112**, 3046–3052 (2002).
- <sup>20</sup>A. Vaas, "Refraction of the direct monotonic sound ray," Technical Report, U.S. Naval Underwater Ordnance Station, Newport, RI (1964).
- <sup>21</sup>M. Porter and H. Buckner, "Gaussian beam tracing for computing ocean acoustic fields," *J. Acoust. Soc. Am.* **82**, 1349–1359 (1987).
- <sup>22</sup>M. Kumar, "World geodetic system 1984: A modern and accurate global reference frame," *Marine Geod.* **12**, 117–126 (1988).
- <sup>23</sup>J. Nelder and R. Mead, "A simplex method for function optimization," *Comput. J.* **7**, 308–313 (1964).
- <sup>24</sup>D. Olsson and L. Nelson, "The Nelder-Mead simplex procedure for function minimization," *Technometrics* **17**, 45–51 (1975).
- <sup>25</sup>J. Stanistreet, D. Risch, and S. Van Parijs, "Passive acoustic tracking of singing humpback whales (*Megaptera novaeangliae*) on a Northwest Atlantic feeding ground," *PLoS One* **8**, e61263 (2013).