Localization of marine mammals near Hawaii using an acoustic propagation model

Christopher O. Tiemann and Michael B. Porter
Center for Ocean Research, Science Applications International Corporation, 10260 Campus Point Dr, San Diego, California 92121

L. Neil Frazer
Department of Geology & Geophysics, University of Hawaii at Manoa, Honolulu, Hawaii 96822

(Received 18 April 2002; accepted for publication 1 December 2003)

Humpback whale songs were recorded on six widely spaced receivers of the Pacific Missile Range Facility (PMRF) hydrophone network near Hawaii during March of 2001. These recordings were used to test a new approach to localizing the whales that exploits the time-difference of arrival (time lag) of their calls as measured between receiver pairs in the PMRF network. The usual technique for estimating source position uses the intersection of hyperbolic curves of constant time lag, but a drawback of this approach is its assumption of a constant wave speed and straight-line propagation to associate acoustic travel time with range. In contrast to hyperbolic fixing, the algorithm described here uses an acoustic propagation model to account for waveguide and multipath effects when estimating travel time from hypothesized source positions. A comparison between predicted and measured time lags forms an ambiguity surface, or visual representation of the most probable whale position in a horizontal plane around the array. This is an important benefit because it allows for automated peak extraction to provide a location estimate. Examples of whale localizations using real and simulated data in algorithms of increasing complexity are provided. [DOI: 10.1121/1.1643368]

PACS numbers: 43.30.Sf, 43.30.Wi [WMC] Pages: 2834–2843

I. INTRODUCTION

Passive acoustic methods of observing marine mammals have been of interest for many years for censusing, behavioral studies, and more recently for ensuring mammals are not present in acoustic ranges during operations which might disturb them.[1–5] The acoustic characteristics of whale songs make them detectable at long ranges using hydrophones.[5–9] For example, low-frequency blue and fin whale sounds can be detected 1600 km away.[10,11] When received over an array of hydrophones, whale songs can be used to estimate a singer's position. Unlike difficult radio tagging, passive acoustic observation methods are unobtrusive; a whale's behavior is unlikely to change because of the observation.[12] Acoustic techniques can observe many individuals at once and are suitable for continuous monitoring applications. In addition, acoustic localization also works when animals are hidden from view, such as at night or when submerged.[4]

This paper will describe a new passive acoustic technique for localizing sound sources based on acoustic propagation modeling with an illustration of the technique to localizing marine mammals using widely spaced receivers. The localization algorithm also provides a novel graphical display of marine mammal location that conveys the confidence of the localization and allows for automatic extraction of location estimates. To demonstrate the benefits of this model-based approach, examples of localizations using both real and simulated data in algorithms of increasing complexity will be provided.

A common technique for localizing marine mammals is that of hyperbolic fixing.[4,5,12–17] The measured difference in arrival time of a whale call recorded on multiple hydrophone pairs produces intersecting hyperbolas indicating the animal's position. When the hydrophone pairs are very closely spaced such as on a short towed array or vertical line array (VLA), hyperbolic fixing is no longer practical. Alternative model-based techniques that exploit either the temporal or spatial structure of the received field are then needed. For instance, the arrival times and amplitudes on a single phone can be used to estimate a whale's range.[9] Alternatively, the interphone phase relationships on a VLA (representing the arrival angles of the multipath) can also be exploited.[18]

The technique described here has several advantages over other localization methods. It uses an acoustic propagation model to account for variations in soundspeed and bathymetry, thus eliminating errors from constant soundspeed and straight-line propagation assumptions inherent to hyperbolic fixing.[16] It can be applied to data from widely spaced individual receivers rather than line arrays. Robustness against environmental variability and acoustic multipath may come from performing some processing in the spectral domain,[12,14] but a formal environmental mismatch study has not been performed. The output of the algorithm is a graphical display that easily conveys mammal location and confidence, and despite the algorithm's added computational complexity, it is suitable for real-time implementation without user interaction.

Acoustic data from the Pacific Missile Range Facility (PMRF) hydrophone network off the western coast of Kauai were used to develop this model-based algorithm.[19] In this data set, the species of interest are humpback whales (*Megaptera novaeangliae*) which are known to congregate

 0001-4966/2004/115(6)/2834/10/$20.00

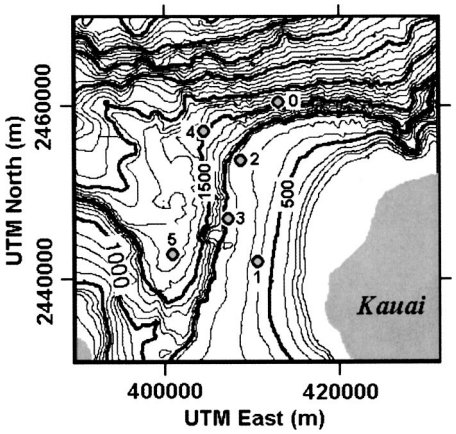

FIG. 1. Bathymetry contours (depths in meters) and hydrophone locations (0–5) at the Pacific Missile Range Facility. Axes are for Universal Transverse Mercator (UTM) Zone 4 coordinates.

near Kauai to breed in winter through spring months after a long migration from the Gulf of Alaska.[20] After describing the available acoustic data set in Sec. II, the localization technique is discussed in Sec. III. Comparisons of localization methods of increasing complexity are presented in Sec. IV.

II. ACOUSTIC DATA

The Pacific Missile Range Facility is an underwater array of over 100 hydrophones in the waters near Kauai, Hawaii. Personnel at PMRF have implemented a near real-time system for transmitting acoustic data from six broadband hydrophones to the Maui High-Performance Computing Center (MHPCC) for analysis. Acoustic data files are posted to MHPCC in 1-min increments. The hydrophones available for use are located 5–20 km apart and are deployed on the sea floor at the locations and depths shown in Fig. 1 and Table I.

Two days of continuous acoustic data from the six hydrophones during March 22 and 23, 2001 were used for algorithm development. Originally sampled at 10 or 20 kHz, the data were low-pass filtered and downsampled to 2 kHz to isolate the frequency band containing most of the energy of the humpback whale songs. Songs are heard on every hydrophone and at all times of day. In many cases, the sounds of multiple marine mammals can be heard simultaneously. While viewing spectrograms of the acoustic data, spectral patterns similar to those associated with humpback whales[21,22] are frequently observed. While it is not practical to listen to every channel of the entire data set duration,

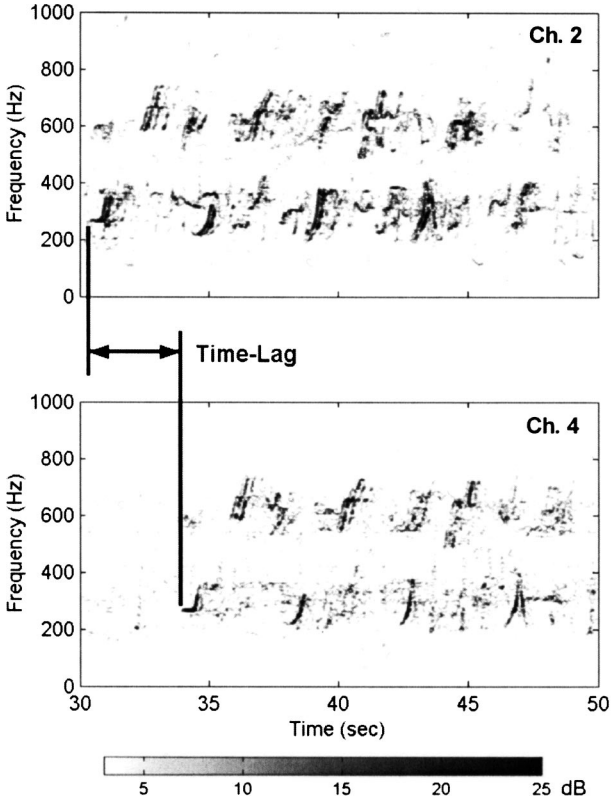

FIG. 2. Spectrograms of acoustic data from hydrophones 2 (top) and 4 (bottom) starting at time 20:16:30 on 3/22/01. A 3.5-s time lag for spectral transients is apparent between the two spectrograms. Spectral patterns resemble those of humpback whale calls.

spectrograms can be examined quickly to confirm that all recordings contained the patterns expected of humpback whale songs.

When spectrograms from all hydrophones for the same time segment are viewed concurrently, similar spectral patterns are often recognized in two or more spectrograms, but offset in time. In such cases, the same whale call is being recorded on multiple receivers, but the time of arrival at the receiver varies according to range from the singer. As an example, Fig. 2 shows spectrograms from hydrophones 2 and 4 for a 20-s segment of data from minute 20:16 on March 22, 2001; the spectrograms were made using 512-point fast Fourier transforms (FFT's) with 90% overlap. A call pattern can be seen repeated on hydrophone 4 approximately 3.5 s after the same pattern on hydrophone 2. It is this difference in arrival times (or time lag) for the same call on two different channels that will be used in the localization process.

TABLE I. Hydrophone positions in geodetic and Universal Transverse Mercator coordinates (Zone 4, WGS84 datum).

Hydrophone No.	Latitude (deg)	Longitude (deg)	UTM North (m)	UTM East (m)	Depth (m)
0	22.246158	−159.842556	2460315.3	413179.3	1638
1	22.080938	−159.867735	2442040.5	410480.1	649
2	22.191175	−159.886739	2454254.8	408590.6	777
3	22.125975	−159.897757	2447044.0	407412.1	843
4	22.215686	−159.929232	2456994.3	404226.6	1560
5	22.091847	−159.957723	2443303.8	401203.5	1768

III. LOCALIZATION ALGORITHM

The model-based localization algorithm consists of two main components: spectral pattern correlation to calculate time lags and ambiguity surface construction to generate a location estimate. Ambiguity surfaces are probabilistic indicators of source location and are constructed through comparison of measured time lags (''data'') to predicted time lags (''replicas''). Replica generation has a hierarchy of increasing modeling complexity, and one goal of this study was to determine how much modeling sophistication is necessary for a correct localization. Example localizations from three techniques using both real and simulated data are presented below to address that question.

A. Spectrogram correlation

Measuring time lags between whale call arrivals at different hydrophones is a critical step in the localization algorithm. The standard method for determining time lags between two signals is through cross correlation, but whether the correlation should be performed on the original wave forms or their spectrograms is open to debate and could depend upon the peculiarities of the signals being processed. Spectrogram correlations are commonly used in whale localization efforts,[4,5,12,14] perhaps because the signal structure remains obvious even in the presence of interferers. Spectrogram correlation may also be more robust than wave form correlation against multipath acoustic arrivals.[12] However, wave forms containing whale calls have been successfully used in both matched-filter[13] and cross correlation[14,16,17] processes. Proponents of wave form approaches argue that the resulting measurements of time lag are more precise.[12,14] Because no formal arguments exist regarding the superiority of a correlation method, both spectral and wave form correlation methods were applied to short segments of the data set in order to determine which method is best for measuring pairwise time lags in the PMRF environment.

The pairwise spectral shape correlations follow an example described by Seem and Rowe.[23] Spectrograms from two hydrophones are digitized, i.e., converted to two levels of intensity (on or off) based on a data-adaptive threshold that guarantees a minimum number of ''on'' pixels per time window. In doing so, the loudest spectral content remains visible in the digitized spectrogram while low-level spectral patterns are hidden, thus adding some robustness against multiple sources. Correlation is done very quickly by performing a logical AND operation on the overlapping region as two digitized spectrograms are shifted past each other. Summing the overlapping pixels provides a correlation score, in units of pixels, whose maximum determines the time lag between channels as well as providing a confidence level of the measurement.

A mathematical description of both the wave form and spectral correlators follows. It assumes that two receivers are separated by a distance d in waters with mean sound speed c. Time series from the two receivers are sampled with a period of Δt and are described by

$$r_i = r(t_i) \quad \text{and} \quad s_i = s(t_i) \quad \text{where} \quad t_i = i \cdot \Delta t. \tag{1}$$

A frame length in seconds, τ_{frame}, is chosen that is slightly longer than a typical whale call from the species of interest. (A 10-s frame length was used with humpback whale calls.) Each frame will contain N samples defined by

$$N = \tau_{\text{frame}} / \Delta t. \tag{2}$$

The mth frame is extracted from each time series:

$$r_i^m = r_{m \cdot N + i} \quad \text{and} \quad s_i^m = s_{m \cdot N + i}, \quad i = 1,...,N. \tag{3}$$

The wave form correlation score at each lag bin l for the mth frame is calculated by

$$c_l^m = \sum_{i=1}^{N} r_i^m \cdot s_{i-l}^m. \tag{4}$$

The lag bin l with the highest wave form correlation score designates the time lag T_w between the two time series for frame m according to

$$T_w^m = l \cdot \Delta t. \tag{5}$$

The spectral correlation is based on the short-time Fourier transform of a time series:

$$R(t, f; \tau_{\text{snap}}) = \int_{-\tau_{\text{snap}}/2}^{\tau_{\text{snap}}/2} r(t-\tau) \cdot e^{-i 2 \pi f \tau} d\tau. \tag{6}$$

Spectrogram R is computed using a FFT to produce a discretized spectrogram. Each spectrogram frame is of length τ_{frame} and has dimensions of N_{freq} frequency bins (256 in this example) and N_{snap} time bins or ''snapshots'' where

$$N_{\text{snap}} = \frac{N}{2 \cdot N_{\text{freq}}} \tag{7}$$

(assume N_{freq} divides N exactly). It follows that the time resolution of the snapshots is

$$\tau_{\text{snap}} = 2 \cdot N_{\text{freq}} \cdot \Delta t \tag{8}$$

and the frequency resolution is

$$\Delta f = \frac{1}{\tau_{\text{snap}}}. \tag{9}$$

Given the notation

$$R_{ij} = R(t_i^{\text{snap}}, f_j) \quad \text{where} \quad i_i^{\text{snap}} = i \cdot \tau_{\text{snap}} \quad \text{and} \quad f_j = j \cdot \Delta f \tag{10}$$

the mth frame is defined as

$$R_{ij}^m = R_{m \cdot N_{\text{snap}} + i, j} \quad \text{where} \quad i = 1,...,N_{\text{snap}} \quad \text{and} \quad j = 1,...,N_{\text{freq}}. \tag{11}$$

A repeat of the above for time series s makes an analogous definition for frame S_{ij}^m. Also calculated is the number of snapshots needed when the maximum possible time lag between sensors is added to the desired frame length:

$$N_{\text{max lag}} = \frac{\tau_{\text{frame}} + d/c}{\tau_{\text{snap}}}. \tag{12}$$

Next, each receiver's spectrogram frame will be digitized or ''pixilated,'' i.e., each time/frequency bin in the spectrogram will be assigned a 0 or 1 according to an adap-

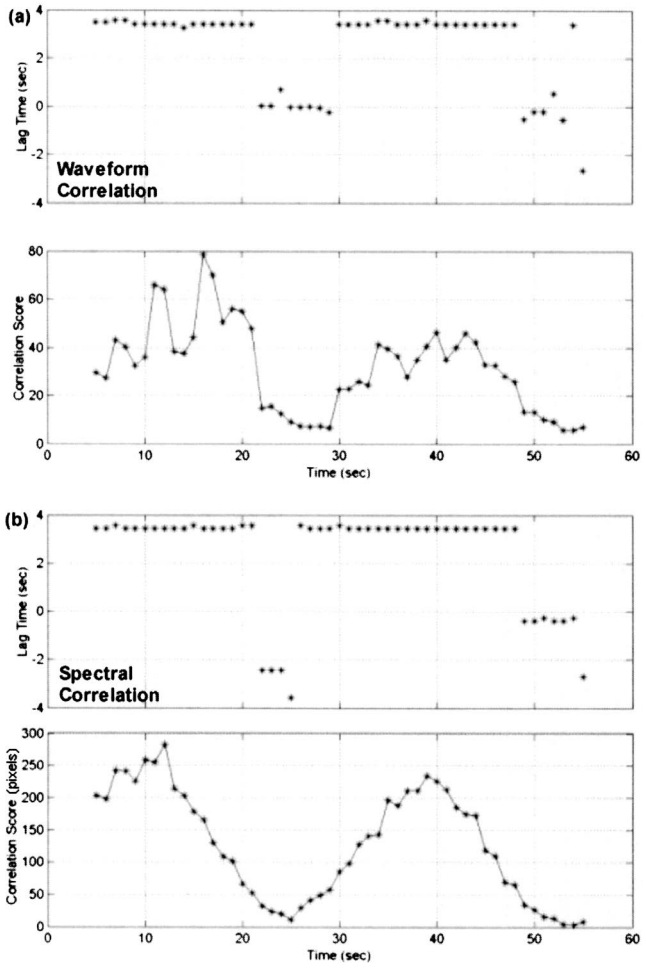

FIG. 3. Time lags and correlation scores output by the cross correlator using wave forms (a) and digitized spectrograms (b). Cross correlations use data from hydrophones 2 and 4 for minute 20:16 on 3/22/01. The 3.5-s time-lag measurement from the correlators agrees with that visually observed in Fig. 2.

tive threshold. A unit step function provides the digitization mechanism:

$$u(x)=\{1 \quad \text{if} \quad x>0$$
$$\{0 \quad \text{if} \quad x\leqslant 0. \tag{13}$$

For each frame R_{ij}^m, a spectral power threshold σ_R^m is calculated that guarantees a minimum number N_{pixels} of "1" pixels (40 per second in this example), such that

$$\sum_i \sum_j u(R_{ij}'''-\sigma_R^m)\geqslant N_{\text{pixels}}. \tag{14}$$

A different threshold σ is calculated for each receiver, and the mth frames are pixelated to make the digitized frames D and E:

$$D_{ij}^m=u(R_{ij}^m-\sigma_R^m) \quad \text{and} \quad E_{ij}^m=u(S_{ij}^m-\sigma_S^m). \tag{15}$$

The spectral correlation score at each lag bin l for the mth frame is calculated by

$$C_l^m=\sum_{i=1}^{N_{\text{snap}}}\sum_{j=1}^{N_{\text{freq}}} D_{ij}^m\cdot E_{i-l,j}^m$$
$$\text{where} \quad l=-N_{\text{max lag}},\ldots,N_{\text{max lag}}. \tag{16}$$

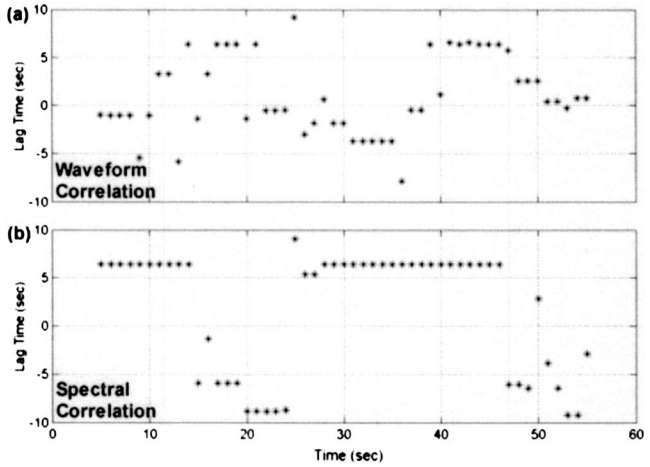

FIG. 4. Time lags output by the cross correlator using wave forms (a) and digitized spectrograms (b). Cross correlations use data from hydrophones 2 and 5 for minute 20:16 on 3/22/01. The consistency of the spectral correlator made it the preferred method for time-lag measurement.

The lag bin l with the highest spectral correlation score designates the time lag T_s between the two mth frames according to

$$T_s^m=l\cdot N_{\text{freq}}\cdot \Delta T. \tag{17}$$

An example of cross correlator output is shown in Fig. 3, where results from both wave form correlation [Fig. 3(a)] and digitized spectral correlation [Fig. 3(b)] are presented for comparison. Data are from hydrophones 2 and 4 for minute 20:16 on March 22, 2001; this time segment includes the data shown in Fig. 2. A time window 10 s long extracts data subsets (frames) to use with each correlation, and the window advances in 1-s increments through the entire minute, calculating a time-lag and correlation score at each step. (Note that correlation scores indicate relative correlation strength among time steps, are in different units, and should not be compared between the two techniques.) In this example, both the wave form and spectrogram correlation methods correctly extract the interchannel time lag of 3.5 s during periods when the whale is singing. Furthermore, the correlator scores drop when the animal stops singing (around 25 s). By setting thresholds on the correlation score, only the most confident of the time-lag estimates are passed to the localization process, thus freeing the correlation output from human examination. A spectrogram correlation score threshold of 100 pixels was used in this processing, and if no correlation score exceeded the threshold for a given time window, no localization was attempted.

Agreement between the two correlation methods is not always as good as that shown in Fig. 3. Typically, the spectral correlator time-lag measurements were more consistent. To illustrate this, Fig. 4 shows output from both the wave form and spectral correlators for the same minute as that in Fig. 3 but for the hydrophone pair 2 and 5. Time-lag measurements from the wave form correlator are quite variable over the minute while the spectral correlation process provides a more stable measurement. Perhaps the scattering in the wave form correlator's output is due to interferers such as other distant animals singing simultaneous songs,[12] but de-

termining the criteria for when one correlation method is better than another remains an interesting area of study which will not be addressed further here. Because of the more consistent measurements provided by the spectral correlator in this environment, it was used in all further analysis to provide time-lag data to the ambiguity surface constructor.

B. Replica generation

The second input needed for ambiguity surface generation is the replica. In this application, the replica is a prediction of time lags that would be measured by every receiver pair combination from a source at every location within a grid of candidate positions around the array. In order to calculate time lags, acoustic travel times from each possible source location to each receiver must be calculated first, but the model complexity used during travel-time calculation can sometimes affect the accuracy of the resulting localization. In efforts to compare the effects of modeling complexity, three replicas were made within a hierarchy of modeling sophistication. Replica computation time increases with added complexity.

Common to all the replicas is the resolution of the candidate source locations. Simulated sources are spaced every 200 m in latitude and longitude in a 30-km square grid around the array. Source frequency is set at 500 Hz, the center of the frequency band of interest, and source depth is assumed to be 10 m, within the range of expected depths for singing humpbacks near Hawaii.[24] While only one source depth is used in the replica generation to follow, the search grid can be expanded to include multiple source depths if needed. Average historical soundspeed profiles for the region were taken from the Generalized Digital Environmental Model (GDEM), and PMRF provided bathymetry data for the range. Geoacoustic properties of the sea floor are those typical of sand:[25] density 1.9 g/cm^3, compressional wave speed 1650 m/s, compressional wave attenuation .8 dB/wavelength. (Values are from Table 1.3 of Ref. 26, based on the work of Hamilton.[27]) The Gaussian beam acoustic propagation model BELLHOP was used to calculate travel times as it can account for depth-dependent sound speed profiles and range-dependent bathymetry.[28,29] Given the small variation in sound speed profiles over the area of study, the assumption of range-independent sound speeds inherent to BELLHOP was acceptable, and refractive effects outside of the plane of propagation were not considered in the modeling.

The simplest replica uses assumptions equivalent to three-dimensional hyperbolic fixing techniques. Sound speed is assumed to be constant at 1510 m/s, and only the direct acoustic path from the shallow source to the true receiver depths determines the acoustic travel time. No bathymetric effects are considered. After travel times from every candidate source position to every receiver position are calculated, the appropriate travel time pairs are subtracted to create a replica matrix of time lags indexed by source position and hydrophone pair.

The next replica in this hierarchy adds a depth-dependent sound speed profile to the acoustic modeling, but bathymetric effects are still ignored; this will be called the ''range-independent'' replica. To illustrate the effect of the downward-refracting sound speed profile on acoustic paths, Fig. 5 shows both the average sound speed profile used in the calculation and the resulting direct acoustic ray paths between the shallow source and hydrophone 0 at several ranges. The mean acoustic sound speed varies with range from the receiver, so travel time will not increase linearly with range. The travel time is still that of the direct acoustic path.

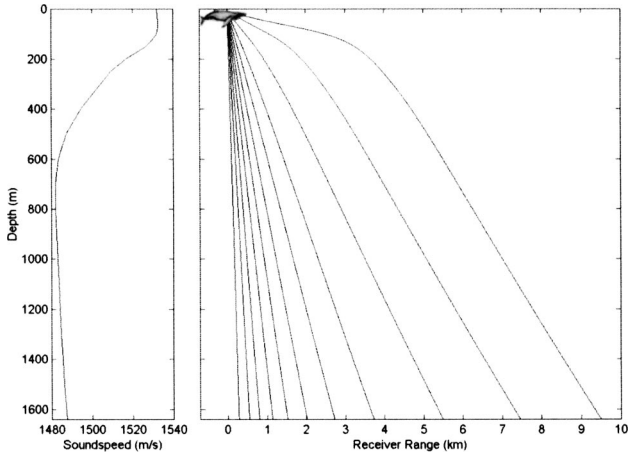

FIG. 5. Average soundspeed profile and predicted direct acoustic ray paths between a shallow whale and hydrophone 0 (1638-m depth) at several ranges. The predicted mean acoustic soundspeed varies with range from the receiver, and no bathymetric effects are included. Travel times from such simulations constitute the ''range-independent'' replica.

Lastly, the ''range-dependent'' replica adds range-dependent bathymetric effects to the acoustic modeling. Note that sound speed profiles are still range independent. The addition of bathymetry contours to the acoustic model allow for multipath arrivals from bottom-reflected paths to be included in the travel time calculation. Figure 6 shows the predicted acoustic ray paths from a shallow source to hydrophones 2 and 5 along two perpendicular bathymetry slices.

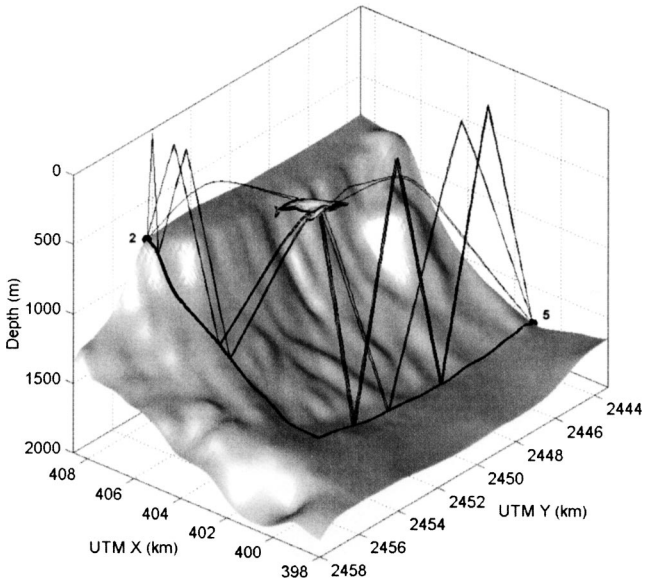

FIG. 6. Predicted direct and reflected acoustic ray paths between a shallow whale and hydrophones 2 and 5 along two perpendicular bathymetry slices. Whale not drawn to scale. The ''range-dependent'' replica allows for both direct and reflected ray paths to be included in the travel time calculation.

While both slices in this example contain both direct and reflected ray paths, in some long-range cases, there is no direct acoustic ray path between a candidate source position and a receiver; only reflected ray paths will connect the two. Because every ray path has a different travel time τ, an average travel time τ_{avg} from all N_{arr} arrivals, weighted by the ray paths' predicted amplitudes a, is used as the single travel time value from a given source location:

$$\tau_{\text{avg}} = \sum_{i=1}^{N_{\text{arr}}} \frac{|a_i| \cdot \tau_i}{a_{\text{sum}}} \quad \text{where} \quad a_{\text{sum}} = \sum_{i=1}^{N_{\text{arr}}} |a_i|. \tag{18}$$

To complete the replica generation process, the replica time lag T_r is made by taking the difference in average travel times for a given receiver pair p from a hypothesized source at search grid position $\overline{x_s} = (x_s, y_s, z_s)$. Like the other replicas, the range-dependent time-lag replica is precalculated for all receiver pair/source position combinations:

$$T_r^p(\overline{x_s}) = \tau_{\text{avg}}^{p_1}(\overline{x_s}) - \tau_{\text{avg}}^{p_2}(\overline{x_s}), \tag{19}$$

where p_1 and p_2 are the two hydrophones making up receiver pair p.

For every replica type, the BELLHOP propagation model also provides an estimate of acoustic intensity $P_n(\overline{x_s})$ for acoustic paths between every source position $\overline{x_s}$ and receiver n. Degradation of the source amplitude, or transmission loss, is saved as part of the replica as well and will be used in scaling the ambiguity surface to follow.

C. Ambiguity surface construction

A singing whale is localized through the construction of an ambiguity surface that is generated in the same way regardless of the type of replica (range independent or range dependent) used. The ambiguity surface is a two-dimensional plan view of the area around the array containing the same latitude/longitude locations as the candidate source positions assumed during replica construction. Although each surface assumes a constant source depth, a different surface can be made for each hypothesized source depth. One input to the localization process is the spectrogram-measured time lags T_s^{mp} and correlation score for frame m and each receiver pair p. Only those measurements with high spectral correlation scores (over 100 pixels) are passed to the localization process. This ensures a high confidence in the resulting location estimates. The replicas of predicted time lags $T_r^p(\overline{x_s})$ and transmission loss $P_n(\overline{x_s})$ serve as another input. Note that replicas are time independent; they only need to be calculated once provided the environment or receiver positions do not change.

For each source position $\overline{x_s}$ and receiver pair combination, the difference between the predicted time lag from the replica and the measured time lag is normalized by the maximum possible time lag between receiver pair p separated by distance d_p. The resulting likelihood scores for each pair form a surface with a minimum where the replica and data agree best. The contours of the surface are accentuated by taking the square root of the likelihood scores to make the new surface L for the mth frame:

$$L_p^m = \sqrt{\frac{|T_r^p(\overline{x_s}) - T_s^{mp}|}{d_p / c}}. \tag{20}$$

In order to represent increased localization uncertainty at long range, the contribution to a localization from a distant receiver pair is diminished; close receiver pairs will contribute more in the ambiguity surface construction. This is done by scaling the likelihood scores by the predicted acoustic intensity. The acoustic transmission loss in dB, α, is calculated for the two acoustic paths from a source to a pair of receivers and summed:

$$\alpha_p(\overline{x_s}) = 20 \log_{10}(P_{p_1}(\overline{x_s}) \cdot P_{p_2}(\overline{x_s})). \tag{21}$$

When an α is found for every candidate source position $\overline{x_s}$, it will form a transmission loss surface of the same dimensions as the likelihood surface L. The likelihood surfaces for all contributing receiver pairs are scaled by their corresponding α and summed to complete construction of the ambiguity surface A for the mth frame:

$$A^m = \sum_p L_p^m \cdot \alpha_p. \tag{22}$$

The surface A represents a planview of the waters around the array for a single source depth, and source location estimates common to many receiver pairs will sum to form a peak indicating the best estimate of source position.

IV. LOCALIZATION COMPARISONS

Because of the added computational complexity required of each level of the replica generation hierarchy, a comparison of localization results from different replicas, plus comparisons to standard hyperbolic fixing techniques, would examine whether any added benefits are worth the additional computational costs. However, that answer depends on the environment under study and the source position. The following sections provide an example in which all localization methods perform equally well plus another example in which only the most sophisticated modeling will produce the correct answer. In the comparisons to follow, time-lag measurements are provided to three localization techniques: two-dimensional hyperbolic fixing, range-independent model-based localization, and range-dependent model-based localization.

A. Real-data localization

The first set of localization comparisons uses time-lag measurements from the same data exhibited in Figs. 2, 3, and 4: recordings of minute 20:16 on March 22, 2001. From one frame of time-lag measurements extracted from that minute, six hydrophone pairs have correlation scores exceeding the score threshold, and their time lags are passed to the localization algorithms. Output from the three localization techniques is presented in Fig. 7 in order of increasing computational complexity. Each frame of Fig. 7 represents a 30-km-square area of ocean around the PMRF array, with hydrophone positions labeled (0–5).

Figure 7(a) represents the traditional technique of plotting intersecting hyperbolic trajectories of possible source

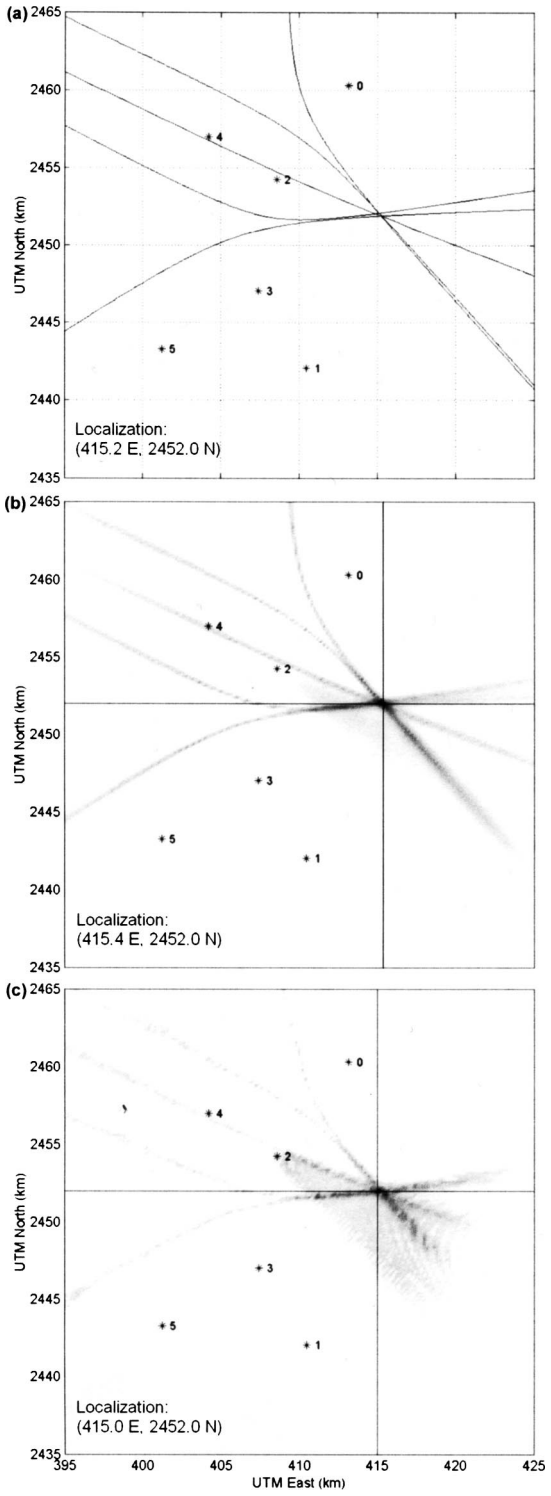

FIG. 7. Plan views of the waters around the PMRF array with hydrophone positions (0–5) indicated. Axes are for UTM Zone 4. Curves from hyperbolic fixing (a) intersect at possible whale positions. Ambiguity surfaces from range-independent (b) and range-dependent (c) model-based localizations indicate whale position estimates with high intensities and crosshairs. Coordinates of the location estimates indicated in figure. Data is from minute 20:16 on 3/22/01.

positions based on time lags. This technique uses no acoustic modeling other than the assumption of a constant soundspeed of 1510 m/s, although the mean horizontal propagation speed versus range from source to hydrophone could vary from 1300 m/s to 1520 m/s depending upon range and depth of the receiver. Note that not all hyperbolic paths intersect precisely at the same point, perhaps due to the errors associated with the constant soundspeed assumption. Nevertheless there is a tight clustering of intersections at approximately 415.2 km east, 2452.0 km north, which is then regarded as the estimate of whale location.

Figures 7(b) and (c) demonstrate the strengths of ambiguity surface visualization. Using the time-lag data and range-independent [7(b)] or range-dependent [7(c)] replicas, ambiguity surfaces were constructed as described in Sec. III C. On these surfaces, areas of peak intensity represent the most confident whale position estimates and are marked with crosshairs. The location estimates from the model-based approaches agree well with each other and the hyperbolic estimate; exact localization coordinates are indicated within each figure frame. Note that ambiguity surfaces still reveal patterns resembling hyperbolas, but the curves have effectively been thickened and stacked in such a way that automatic identification of the most probable source location is possible. The jaggedness of the range-dependent curves is due to variability in travel time and transmission loss predictions caused by bathymetry effects. Furthermore, the narrowness of the ambiguity surface peak convey high confidence in the localization. A sharp peak implies that many receiver paris had the same location estimate in common; a broad peak suggests greater uncertainty in the localization as several pairs' location estimates failed to overlap at a common point.

In this localization example, all techniques agree well, and localizations are in close proximity regardless of replica complexity. The agreement between all techniques is probably due to the relatively short ranges from source to receiver and deepness of the water; direct acoustic paths to all receivers exist, so accounting for bathymetric effects is not necessary for a correct answer. Unfortunately, no independent visual surveys are available during the times of the acoustic recordings, so location estimates cannot be verified through other means.

The analysis described here was applied to many other short time segments throughout the two days of acoustic data. Localization using the constant-soundspeed replica, not shown in the comparisons above, was included as well. In every case, a source was confidently localized by the model-based techniques through a contribution of four or more receiver pairs. The acoustic data from those times were then played back to verify the presence of a marine mammal. However, when using hyperbolic fixing methods, the tight grouping of intersections like those in the example above was not always seen, sometimes making source location difficult to determine. It was hypothesized that any advantages of the most sophisticated range-dependent replica over the other replicas would best be seen in localizations of sources at long range from the receivers. This could best be tested by placing a simulated whale at the extent of the search grid.

B. Simulated localization

To demonstrate a situation when the full complexity of the range-dependent model-based replica is necessary for a correct localization, a simulated source is placed in the

southwest corner of the search grid very near the border (396 km east, 2436 km north, 10 m depth) and the acoustic model BELLHOP used to simulate travel times from the source to all receivers. The difference in simulated travel times became the simulated time-lag data passed to the localization algorithms.

Figure 8 shows the resulting source location estimates from three techniques. The hyperbolic fixing method, shown in Fig. 8(a), has many hyperbola intersections, each indicating a possible source location. Because the intersections are scattered over several square kilometers, determining a single location estimate is difficult. The range-independent model output, shown in Fig. 8(b), shows some increases in intensity on the ambiguity surface indicating likely source positions. However, the contributions from individual receiver pairs do not stack up correctly to form a single peak at the true source location. Instead, a coincidental intersection of ambiguity surface curves puts a peak 15.3 km away from the true source location. The ambiguity surface from the range-dependent model, Fig. 8(c), correctly identifies the true source position, but it is expected to do so since the data and replica in this test will have perfect agreement at the source location. This one simulation illustrates that at ranges where refractive and bathymetric effects are important the assumptions inherent to hyperbolic fixing and the range-independent replica break down, leading to an incorrect localization.

In efforts to quantitatively compare the localization errors of the different techniques, a simulation like the one above was repeated many times while moving the simulated source through every search grid position around the array. The distance between the resulting location estimate and the source was recorded for each source position. The localization errors were then assembled to make an error map like those of Fig. 9. The three replicas of increasing complexity as described in Sec. III B were used in the localization process, each generating its own error map. Figure 9(a) shows the error map for the simplest replica which uses assumptions equivalent to hyperbolic fixing techniques: constant soundspeed and straight-line, direct acoustic paths with no bathymetric effects. Figure 9(b) shows the error map when the range-independent replica is used in the localization. Note that an error map for the range-dependent replica is not shown because it always correctly identifies the source location; the simulated data and replica are identical.

Of interest in Fig. 9 is that localization is almost perfect regardless of replica type when the source is close enough to the receivers for a direct acoustic path to exist. In cases where bathymetric effects can be ignored, even the assumptions of the simplest hyperbolic fixing method are still suitable for a correct localization. The advantages of using the full range-dependent replica are apparent at the outer limits of the search grid where localization errors from using simple models can be as high as 25 km. Only the range-dependent replica that accounts for bottom interactions correctly locates the source. Therefore, when trying to extend target localization far beyond array boundaries, one must balance the increased accuracy of the more sophisticated replica model against the increased costs of longer calculation time and higher environmental characterization requirements.

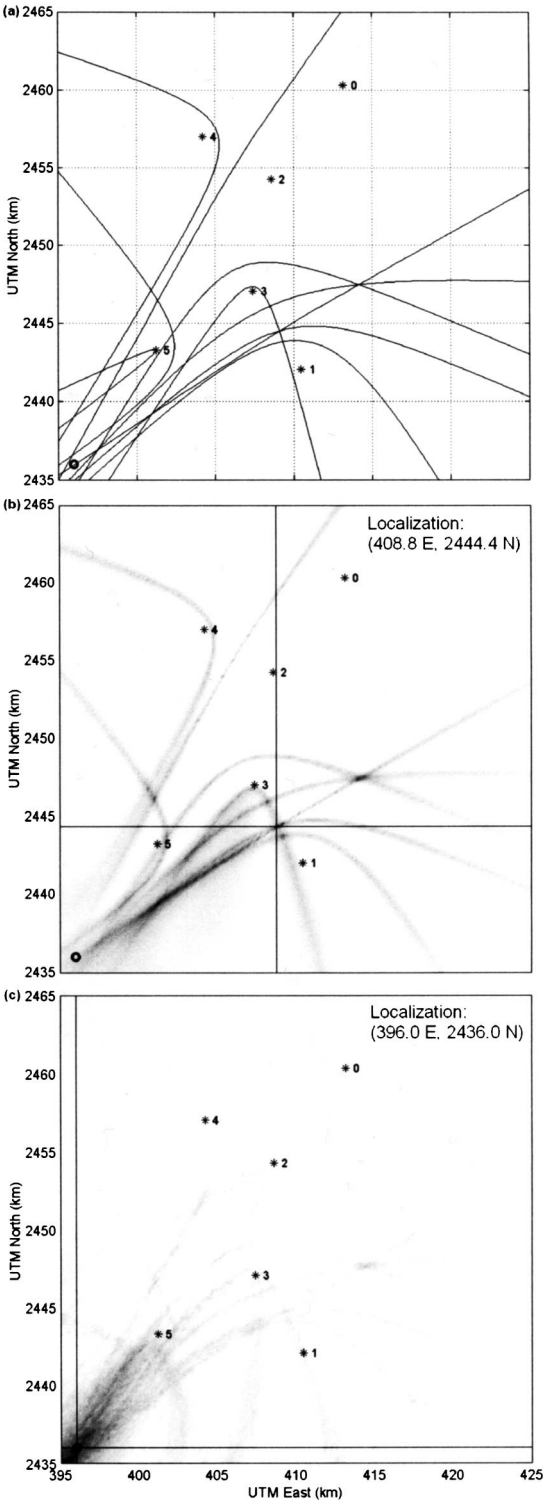

FIG. 8. Plan views of the waters around the PMRF array with hydrophone positions (0–5) indicated. Axes are for UTM Zone 4. Curves from hyperbolic fixing (a) intersect at many possible whale positions. Ambiguity surfaces from range-independent (b) and range-dependent (c) model-based localizations indicate whale position estimates with high intensities and crosshairs. Data are from a simulated whale at 396 km east, 2436 km north with position indicated by circle. Coordinates of the model-based location estimates indicated in figure.

The error maps of Fig. 9 are an example of how one can quantify errors associated with the different techniques

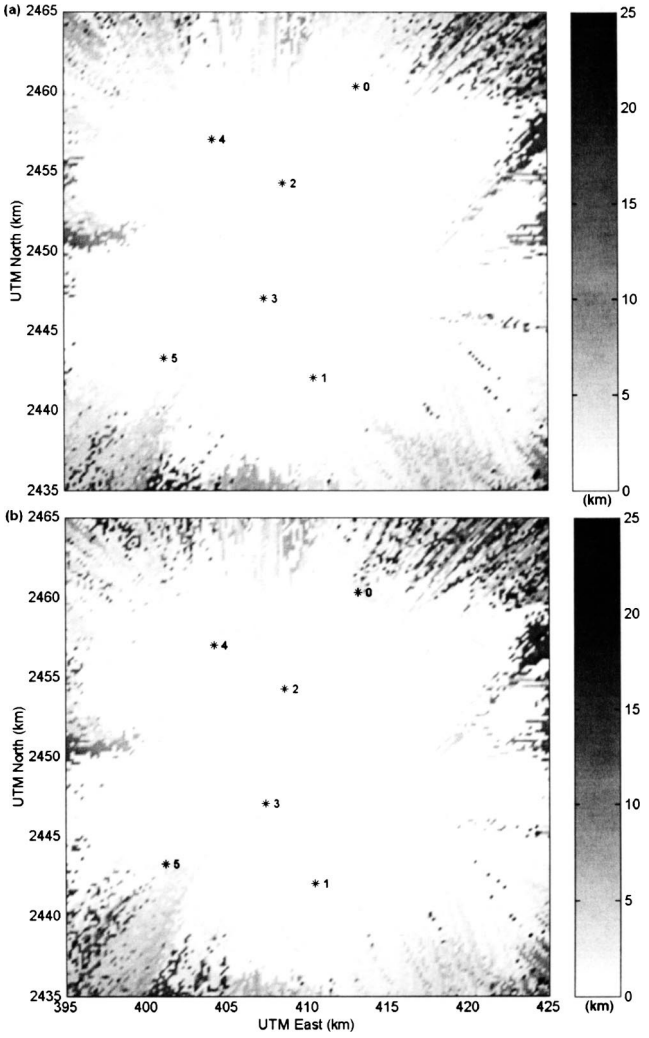

FIG. 9. Maps of localization error on plan views of the waters around the PMRF array with hydrophone positions (0–5) indicated. Axes are for UTM Zone 4. Maps indicate distance in kilometers between source and location estimate for a simulated source at the map coordinates. Results from simple hyperbolic replica shown in (a); range-independent replica shown in (b). Errors increase when far outside array as bathymetric effects become more important with increased range.

through simulation, but to truly measure the error an experiment localizing a controlled source of known location is required. However, any error measurements, real or simulated, will be specific to the environment and receiver geometries used and thus are not easily generalized. It is also difficult to get formal bounds on uncertainties in the localization, such as those resulting from environmental mismatch or the correlation process for example. One strategy for measuring uncertainty could involve adding mismatch in a Monte Carlo fashion to the environment used in replica and simulated data generation and then repeating the localization process.

V. DISCUSSION

The purpose of this work is to introduce a new passive acoustic technique, with advantages over traditional methods, for localizing singing marine mammals, humpback whales in particular. Based on acoustic propagation modeling, it claims increased accuracy in geometries where acoustic bottom interaction becomes important. The algorithm also provides a visual display of whale location that is easy to interpret and allows for automatic location extraction. While the few localization examples shown here, plus many others not presented, build confidence in the algorithm, it is recognized that the algorithm has yet to be verified with other independent methods such as by visual observation or controlled source localization.

One question that arises is what amount of acoustic modeling complexity is really necessary for localizations of a desired accuracy. Each level of the modeling hierarchy has its advantages and disadvantages. For example, the range-independent replica used in this work can be quickly calculated in minutes and requires no prior knowledge of an environment's bathymetry or geoacoustic properties. The range-dependent replica can improve localization accuracy, particularly at long ranges, but it requires 100 times more computation time. The advantages of the range-dependent replica may drastically increase in areas with complicated bathymetry or in shallow water. Ultimately, the choice of model lies with the user to balance environmental definition and replica precalculation time versus localization accuracy and range, and even traditional hyperbolic fixing methods should remain an option in some geometries.

It should be stressed that although replica precalculation can be a several-hour process, this step needs to be repeated only as often as the environment or array geometry changes; the remaining spectrogram correlation and localization calculations are relatively simple. In the analysis of data from PMRF, the localization could be completed within the data update period of 1 min. The algorithm can run without operator oversight by requiring high localization scores, such as over 75% of maximum possible score, be met before declaring a localization in order to minimize false alarms. It is also rapid enough for near real-time processing. Both of these qualities make it a good candidate for continuous, long-term monitoring of marine mammal activity, provided the animals are vocalizing.

To demonstrate how this algorithm could be a tool for behavioral studies, Fig. 10 shows the most confident whale location estimates from the 24 hours of March 23, 2001. The locations of ambiguity surface peaks that were over 75% of the maximum score are shown as points on this plan view. From this plot, one could conclude that on this day singing whales preferred to stay near the shore of Kauai instead of venturing out into deeper water. Through acoustic studies over longer time periods and ranges, common travel routes may become apparent, especially when used in conjunction with other complementary techniques such as visual observation and tagging.

The algorithm presented here has a modular design that adds to its flexibility and facilitates advancement. For example, should wave form cross correlations offer advantages over spectrogram correlations, it is easy to substitute that step in the processing. Because replica generation is independent of the visualization process, it does not have to be limited to ray theory; full wave acoustic models may easily be substituted for still further improvements in accuracy. Lastly, while the algorithm described here was used in a two-

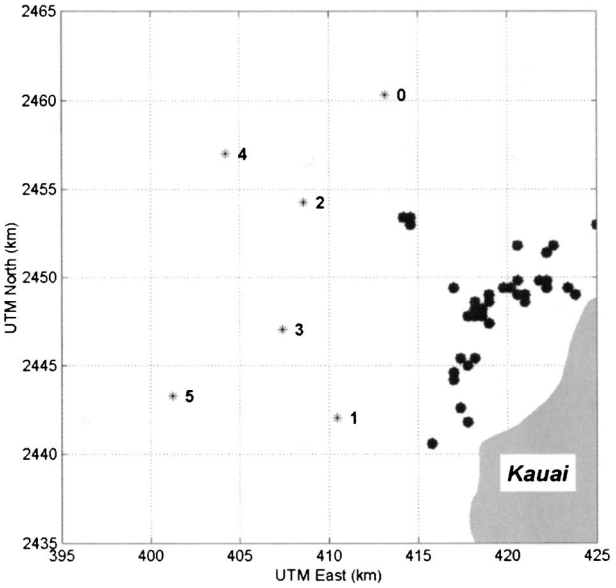

FIG. 10. Plan view of the waters around the PMRF array with hydrophone positions (0–5) indicated. Axes are for UTM Zone 4. Points show locations of the most confident whale location estimates during the 24 hours of 3/23/01. Such continuous, long term monitoring can provide clues to animal behavior.

dimensional search (latitude/longitude) it has an immediate generalization to a full three-dimensional volumetric search.

ACKNOWLEDGMENTS

We gratefully acknowledge the support of the Pacific Missile Range Facility (Jim Hager) and the Maui High-Performance Computing Center (Robert Desonia and D. J. Fabozzi) in establishing the real-time monitoring facility. Several colleagues at SAIC were also instrumental in this work. In particular, thanks to Dave Gever for implementing the web-based control software and Richard Bachman for environmental characterization of the site. Thanks to Paul Hursky for preparing all the raw acoustic data for use. Finally, we thank Herb Freese for general discussions. This work was supported by the National Defense Center of Excellence for Research in Ocean Sciences (CEROS Contract No. 47316). Additional support was provided by the Office of Naval Research (ONR Contract No. N00014-00-D-0115).

[1] H. E. Winn, R. K. Edel, and A. G. Taruski, ''Population estimate of the humpback whale (*Megaptera novaeangliae*) in the West Indies by visual and acoustic techniques,'' J. Fish. Res. Board Can. **32**, 499–506 (1975).
[2] W. A. Watkins, ''Activities and underwater sounds of fin whales,'' Sci. Rep. Whales Res. Inst. **33**, 83–117 (1981).
[3] P. O. Thompson, W. C. Cummings, and S. J. Ha, ''Sounds, source levels, and associated behavior of humpback whales, Southeast Alaska,'' J. Acoust. Soc. Am. **80**, 735–740 (1986).
[4] C. W. Clark, W. T. Ellison, and K. Beeman, ''Acoustic tracking of migrating bowhead whales,'' IEEE Oceans 1986 Conference Proceedings, pp. 341–346.
[5] A. S. Frankel, C. W. Clark, L. M. Herman, and C. M. Gabriele, ''Spatial distribution, habitat utilization, and social interactions of humpback whales, *Megaptera novaeangliae*, off Hawai'i determined using acoustic and visual techniques,'' Can. J. Zool. **73**, 1134–1146 (1995).
[6] T. F. Norris, M. McDonald, and J. Barlow, ''Acoustic detections of singing humpback whales (*Megaptera novaeangliae*) in the eastern North Pacific during their northbound migration,'' J. Acoust. Soc. Am. **106**, 506–514 (1999).
[7] J. A. Thomas, S. R. Fisher, and L. M. Ferm, ''Acoustic detection of cetaceans using a towed array of hydrophones,'' Rep. Int. Whal. Commn. **8**, 139–148 (1986).
[8] H. E. Winn and L. K. Winn, ''The song of the humpback whale (*Megaptera novaeangliae*) in the West Indies,'' Mar. Biol. (Berlin) **47**, 97–114 (1978).
[9] M. A. McDonald, J. A. Hildebrand, and S. C. Webb, ''Blue and fin whales observed on a seafloor array in the Northeast Pacific,'' J. Acoust. Soc. Am. **98**, 712–721 (1995).
[10] C. W. Clark, ''Application of the U.S. Navy underwater hydrophone arrays for scientific research on whales,'' Sci. Rep. Int. Whal. Comm. **44**, 1–12 (1994).
[11] C. W. Clark, ''Acoustic behaviors of blue and fin whales,'' paper presented at the 24th International Ethology Conference, Honolulu, HI, 1995.
[12] V. M. Janik, S. M. Van Parijs, and P. M. Thompson, ''A two-dimensional acoustic localization system for marine mammals,'' Marine Mammal Sci. **16**, 437–447 (2000).
[13] 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).
[14] C. W. Clark and W. T. Ellison, ''Calibration and comparison of 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).
[15] F. Desharnais, M. Laurinolli, A. Hay, and J. A. Theriault, ''A scenario for right whale detection in the Bay of Fundy,'' MTS/IEEE Oceans 2000 Conference Proceedings, Providence, Rhode Island, September 11–14, 2000 (Holland Publications, Escondido, California, 2000), pp. 1735–1741.
[16] J. L. Spiesberger and K. M. Fristrup, ''Passive localization of calling animals and sensing of their acoustic environment using acoustic tomography,'' Am. Nat. **135**, 107–153 (1990).
[17] S. Mitchell and J. Bower, ''Localization of animal calls via hyperbolic methods,'' J. Acoust. Soc. Am. **97**, 3352–3353 (1995).
[18] 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).
[19] C. O. Tiemann, M. B. Porter, and L. Neil Frazer, ''Automated model-based localization of marine mammals near Hawaii,'' MTS/IEEE Oceans 2001 Conference Proceedings, Honolulu, Hawaii, November 5–8, 2001 (Holland Publications, Escondido, California, 2001), pp. 1395–1400.
[20] C. S. Baker, L. M. Herman, A. Perry, W. S. Lawton, J. M. Straley, A. A. Wolman, G. D. Kaufman, H. E. Winn, J. D. Hall, J. M. Reinke, and J. Ostman, ''Migratory movement and population structure of humpback whales (*Megaptera novaeangliae*) in the central and eastern North Pacific,'' Mar. Ecol.: Prog. Ser. **31**, 105–119 (1986).
[21] P. O. Thompson and W. A. Friedl, ''A long term study of low frequency sounds from several species of whale off Oahu, Hawaii,'' Cetology **45**, 1–19 (1982).
[22] D. J. McSweeney, K. C. Chu, W. F. Dolphin, and L. N. Guinee, ''North Pacific humpback whale songs: A comparison of southeast Alaskan feeding ground songs with Hawaiian wintering ground songs,'' Marine Mammal Sci. **5**, 139–148 (1989).
[23] D. A. Seem and N. C. Rowe, ''Shape correlation of low-frequency underwater sounds,'' J. Acoust. Soc. Am. **95**, 2099–2103 (1994).
[24] J. D. Darling, ''Migration, abundance, and behavior of Hawaiian humpback whales (*Megaptera novaeangliae*),'' Ph.D. dissertation, University of California, Santa Cruz, 1983.
[25] P. J. Fischer and D. R. Thor, ''A preliminary oceanographic-environmental evaluation of the BSURE area, northwest of Kauai, Hawaii,'' California State Univ. Northridge report CSUN-MS-75-1, p. 31 (1975).
[26] F. B. Jensen, W. A. Kuperman, M. B. Porter, and H. Schmidt, *Computational Ocean Acoustics* (American Institute of Physics, Woodbury, NY, 1999), p. 41.
[27] E. L. Hamilton, ''Geoacoustic modeling of the sea floor,'' J. Acoust. Soc. Am. **68**, 1313–1340 (1980).
[28] M. B. Porter, ''The KRAKEN normal mode program,'' SACLANT Undersea Research Centre Memorandum (SM-245)/Naval Research Laboratory Mem. Rep. 6920 (1991).
[29] M. B. Porter and Y. C. Liu, ''Finite-Element Ray Tracing,'' Proceedings of the International Conference on Theoretical and Computational Acoustics, Eds. D. Lee and M. H. Schultz (World Scientific, Singapore 1994), pp. 947–956.