PAPER

Marine Mammal Monitoring on Navy Ranges (M3R): A Toolset for Automated Detection, Localization, and Monitoring of Marine Mammals in Open Ocean Environments

AUTHORS

Susan M. Jarvis
Ronald P. Morrissey
David J. Moretti
Nancy A. DiMarzio
Jessica A. Shaffer
Naval Undersea Warfare Center Division, Newport, RI

ABSTRACT

Navy sonar has been associated with a number of marine mammal stranding events worldwide. As a result, determining the effects of anthropogenic noise on marine mammals is currently an active area of research. The development of methods to detect and localize the animals in their native environments is key to advancing this research and our understanding. This paper presents a collection of algorithms for automated passive acoustic detection, classification, and localization of vocalizing marine mammals in open ocean environments. The tool set known as M3R (Marine Mammal Monitoring on Navy Ranges) uses the large fields of wide-bandwidth bottom-mounted hydrophones that are part of the U.S. Navy's undersea ranges to listen for vocalizing whales. M3R employs time-frequency analysis to passively detect whale vocalizations; it then aligns detections among neighboring hydrophones to determine the difference in times of arrival (TDOA) of each vocalization. Sets of TDOA are then used to determine 2-D or 3-D position points using hyperbolic localization techniques. An M3R system is capable of continuous, automated, real-time monitoring of over 200 wide-bandwidth hydrophones covering 2,000+ km^2 of open ocean. M3R is typically used in a collaborative fashion to localize animals in support of tagging exercises, visual surveys, and behavioral response experiments.

Keywords: passive acoustic monitoring, marine mammal monitoring, detection and localization

Introduction

Navy mid-frequency sonar has been associated with several mass strandings of marine mammals (e.g., Cox et al., 2006; England, 2001; Southall et al., 2005). As a consequence, both the scientific community and the wider public are questioning the impact of active sonar emissions on marine mammals. Determining the effects of anthropogenic noise on marine mammals requires the ability to monitor the animals in their native environments. This in turn requires the ability to reliably detect and localize and, in many cases, count the animals.

Passive acoustics has become a popular modality for studying marine mammals and fish in open ocean environments. Passive acoustic monitoring (PAM) methods employ a variety of sensors including single hydrophones hand-deployed from the surface (Perkins, 1987; Watkins, n.d.), deployed autonomous recording devices (Clark, 2012; Lammers et al., 2008; Wiggins & Hildebrand, 2007), towed arrays (Pavan et al., 2010), and sea gliders (Klink et al., 2012). Many PAM systems record data for post-processing, while others employ *in situ* processing to detect, classify, and/or localize marine mammals.

Other methods for studying marine mammals and fish are largely visual. They include ship and aerial line-transect surveys, capture experiments (fish), captive animals, and dive experiments. PAM offers the potential for larger areas of coverage and longer-term observations than visual methods. It also allows for subsurface "observation" and is unaffected by darkness and typically less affected by weather. PAM also has its drawbacks. The animal subjects under study must vocalize, and there can be uncertainty about what one is listening to without corresponding visual observations or tagging of individuals with data recording tags.

Navy undersea ranges consist of large fields of widely spaced, broadband hydrophones. Since 1999, these range hydrophones have been used as

sensors of opportunity to record and monitor vocalizations from marine mammals. The M3R (Marine Mammal Monitoring on Navy Ranges) system is a collection of real-time detection, analysis, and display tools that have been developed over time to automatically process multiple acoustic data streams specifically from undersea ranges. Working with teams of vessel-based researchers, vocalizations detected by M3R have been definitively associated with specific species through visual observation and tagging of individual animals. The catalog of M3R detection, localization, and classification data when viewed over time paints a picture of vocal activity across large (1,200–2,000+ km^2) open ocean areas. These data are actively used by a number of researchers to infer other information from response to active sonar (McCarthy et al., 2011) to animal abundance (Marques et al., 2009; Moretti et al., 2010) to population health (Moretti, 2011).

Methods

M3R processing consists of the following functional blocks: data acquisition, call detection, data association, position determination, real-time display, and postprocessing analysis. All of these functional blocks are fully automated. The system is implemented using a cluster of commodity computers that are connected via two 1-Gbit networks and one 100-Mbit network. The two 1-Gbit networks are private to the cluster and are not available to any wider networks. These networks are dedicated to the transport of data among the processing nodes in the cluster. The 100-Mbit network is public, and access to it can be granted to other local area or wide

FIGURE 1

Block diagram of the architecture of the M3R system.

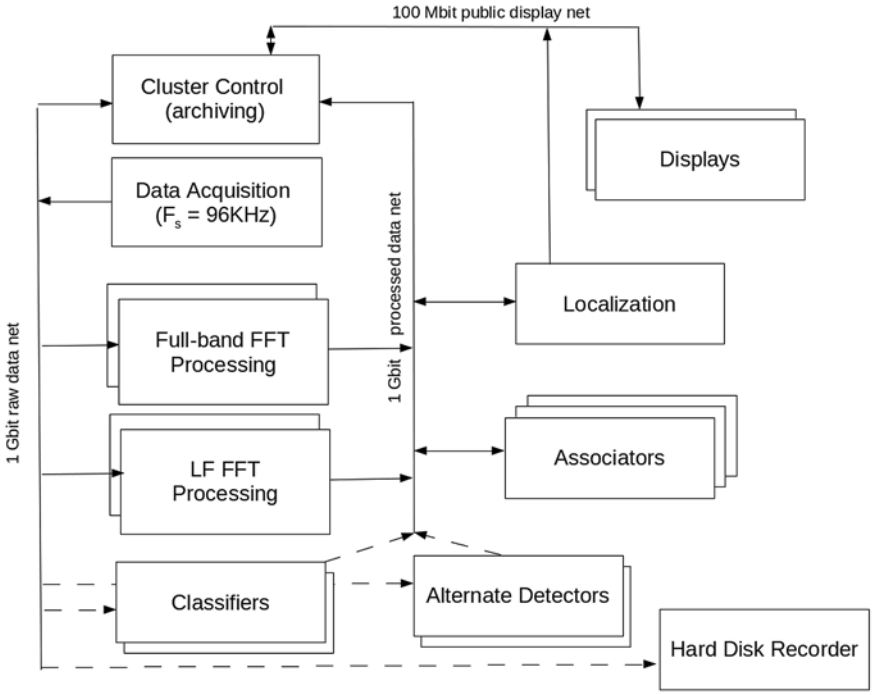
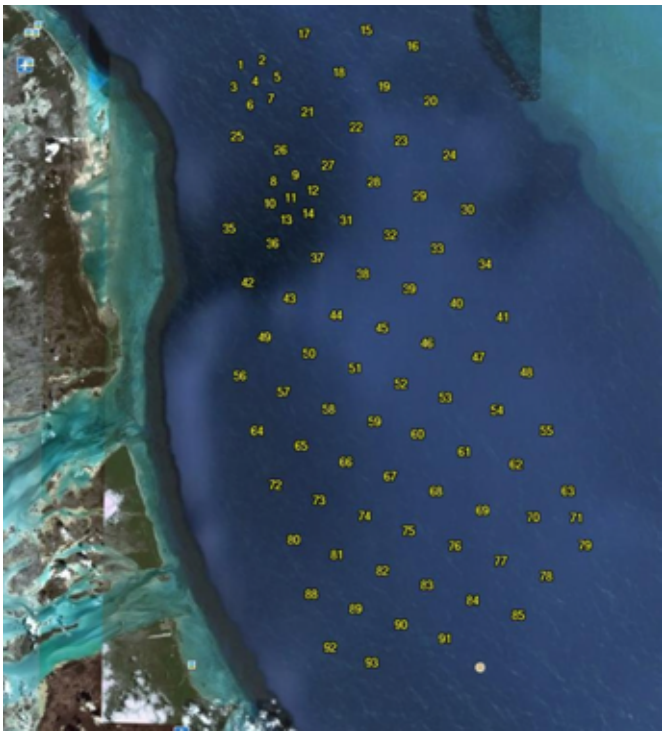


FIGURE 2

Layout of the bottom-mounted range hydrophones at AUTEC. Ninety-three wide-band hydrophones are spaced approximately 4 km apart covering an area of approximately 1,000 km^2 in the Tongue of the Ocean, Bahamas.



area networks. Data for display are made available on the 100-Mbit network. The various functional blocks operate in parallel and are each implemented individually on subsets of cluster nodes (Figure 1).

Description of Sensor Fields and Data Acquisition

The M3R system is currently installed at three of the U.S. Navy's major undersea ranges, the Atlantic Undersea Test and Evaluation Center (AUTEC), off Andros Island, Bahamas; the Southern California Offshore RangE complex (SCORE), off San Clemente Island, California; and the Pacific Missile Range Facility (PMRF), off Kauai, Hawaii. These undersea ranges each contain fields of between 93 and 219 wide-bandwidth hydrophones. At both AUTEC and SCORE, the hydrophones are spaced approximately 4 km apart and are located in deep water with sensor depths averaging approximately 1,600–1,800 m (Figure 2). At PMRF, hydrophone spacing varies from approximately 1 to 7 km, depending on water depth, which varies from less than 80 m to more than 4,700 m. The hydrophones' beam patterns are nearly hemispherical toward the surface, and the acoustic output from each hydrophone is individually available to M3R on shore. The bandwidth of most range hydrophones is 50 Hz to 50 kHz, but a subset of the sensors at all three sites is high-pass filtered at 8 kHz by the in-water electronics. Additionally, 18 of the deep-water hydrophones at PMRF have an upper frequency cutoff of 20 kHz. All hydrophone data are digitized synchronously with a master time code at a sampling frequency of 96 kHz and 16-bit resolution. Time-tagged, digitized raw data are then multicast over one of the private Gbit networks for processing. The aggregate raw data rate across an entire range is substantial. The 93 hydrophones at AUTEC generate approximately 1 GB of raw data per minute, while the 178 hydrophones at SCORE generate just over 2 GB per minute and the 219 hydrophones at PMRF produce more than 2.5 GB per minute.

Call Detection

The first challenge in call detection was characterization of the signals of interest. Early audio-band (F_s = 48 kHz) recording of hydrophones at AUTEC indicated that the broadband clicks from sperm whales were regularly present (Ward et al., 2000). Those recordings also contained narrow-band, frequency swept signals (i.e., whistles) and other broadband clicks. These whistles and clicks were presumed to be from dolphins. Based on these data, the M3R detector was initially designed to simply characterize marine mammal vocalizations as either clicks (broadband events) or whistles (narrow-band event). This was done using spectrogram analysis based on a frequency-domain energy detector (Morrissey et al., 2006). Over time, more information on the vocalizations produced by different species has become available as has the ability to process wider bandwidths. M3R continues to use spectrogram analysis as its core detection algorithm because of the flexibility and generality it offers in detecting many types of vocalizations.

The core M3R system processes hydrophone data in two ways, full bandwidth processing and low-frequency (LF) processing. In full bandwidth processing, a spectrogram $X_i(f,n)$ of the time series data from each of the hydrophones (i = 1 to the number of range hydrophones) is formed using 2,048-point fast Fourier transforms (FFTs) with a rectangular window and 50% overlap. With F_s = 96 kHz, this results in a frequency resolution of 46.875 Hz and a time step of 10.67 ms. Each time-frequency bin of $X_i(f,n)$ is compared to a time-varying threshold $D_i(f,n)$. The threshold is set to be a multiplicative factor k above an exponential average of the power $N_i(f,n)$ within frequency bin f.

$$N_i(f,n) = (1-\alpha)X_i(f,n) + \alpha N_i(f,n-1)$$

where $0 < \alpha < 1$ and

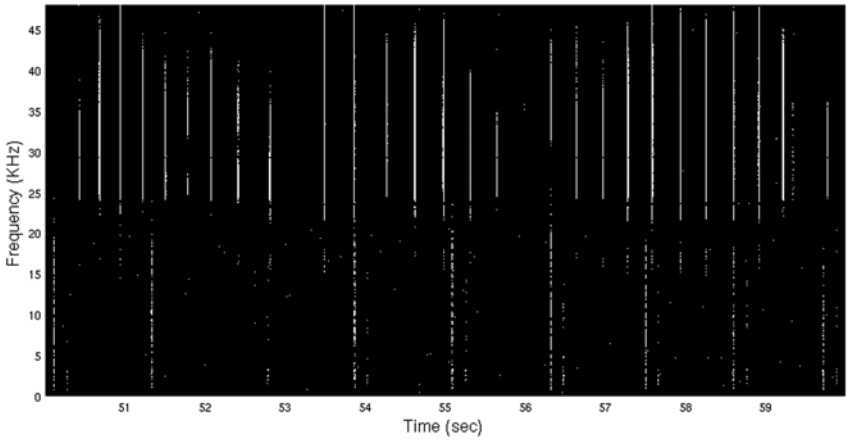
$$D_i(f,n) = kN_i(f,n).$$

The output of the FFT detector for each hydrophone is a binary valued detection spectrogram $Q_i(f,n)$, which contains a 1 in each time-frequency bin that exceeds $D_i(f,n)$ and a 0 everywhere else (Figure 3). The parameter α has been empirically chosen to provide an averaging time constant of 0.2 s. Threshold factor k has also been empirically set. The current setting has a measured false alarm rate of approximately 20 false alarms per second (Ward et al., 2008) and a theoretical probability of false alarm, P_{fa} = 0.214 (Ward et al., 2011). When viewed on a detection spectrogram (Figures 3 and 14), these false alarms from the FFT detector appear as speckle.

The time-frequency resolution of the full-band processing is matched to the detection of sperm whale clicks, which are approximately 10-ms long (Møhl et al., 2003; Ward, 2002), and to delphinid whistles, which sweep over several kHz (Au, 1993). It is less well matched to the detection of impulsive dolphin clicks and short-duration beaked whale clicks. However, there are practical data bandwidth

FIGURE 3

Full bandwidth detection spectrogram showing clicks characteristic of Blainville's beaked whale (*Mesoplodon densirostris*) in the band above 24 kHz over clicks characteristic of a (distant) sperm whale (*Physeter macrocephalus*) in the 1- to 20-kHz region. The speckle is caused by false alarms from the FFT detector.



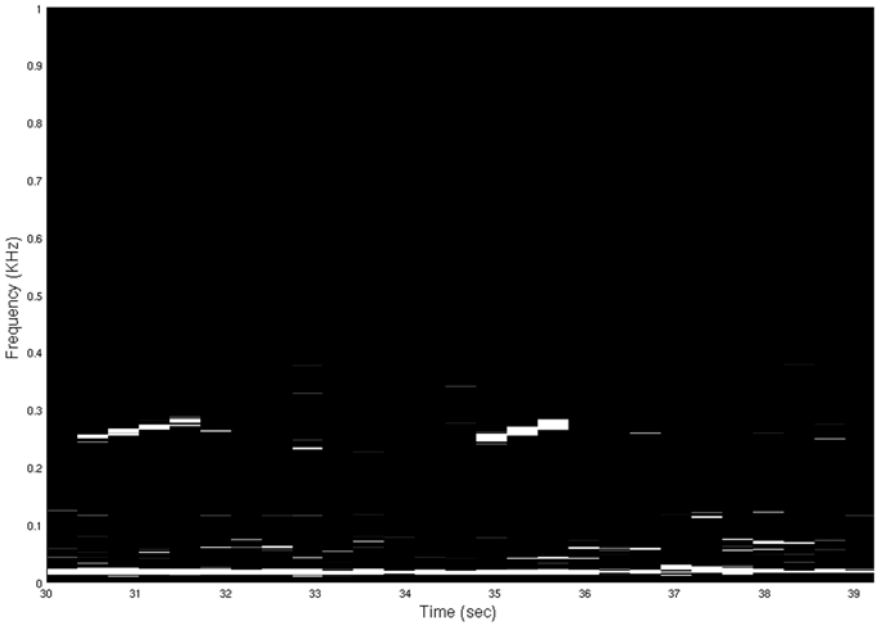
and processing speed limitations that prevent increasing either the time or frequency resolution of the full-band FFT processing. Also, M3R is intended to be a general system. The full-band resolutions of 46.875 Hz and 10.67 ms, while admittedly sub-optimal for some species, allow detection of a wide variety of odontocete vocalizations.

The full-band FFT detector is not well suited to the detection of calls from baleen whales, however. Several baleens like blue, fin, humpback, and Sei whales emit very LF (<100 Hz) narrow-band calls (Baumgartner et al., 2008; Rankin & Barlow, 2007; Thompson et al., 1986), while humpbacks and minkes, for example, can produce narrow-band calls in the 0.5- to 3-kHz range (Martin et al., 2013). The full-band FFT detector does not have the frequency resolution required to process these calls. The LF FFT detector was added to M3R in 2010 specifically to address this deficiency. The LF detector only processes data from those hydrophones without the 8-kHz high-pass filter. It decimates the time series data $x_i(n)$ by 16 to achieve a 6-kHz sampling frequency and then forms a 4,096-point FFT with rectangular window and 50% overlap. This results in 1.46-Hz frequency resolution and 341.3-ms time resolution in the band from 0 to 3 kHz. Again, each time-frequency bin of the spectrogram is compared to a time-varying threshold creating a binary detection spectrogram $Q_{LFi}(f,n)$ (Figure 4).

In an M3R computer cluster, two to four nodes are dedicated to FFT processing. The output of the M3R FFT detectors are bundled into detection reports that are multicast over the second Gbit network for display and further processing. A detection report is generated for each FFT from each hydrophone. The detection reports contain the time of the start of the FFT window, the hydrophone number, and the binary thresholded spectral data packed into 32-bit words. Other statistics including sampling frequency, FFT size, peak frequency, and number of bins above threshold are also included in the report. By hard-limiting the FFT data and reporting only the detection spectra, a data bandwidth

FIGURE 4

LF detection spectrogram showing up-sweep vocalizations determined from aural analysis to be from a humpback whale (*Megaptera novaeangliae*) and an LF (~20 Hz) tone of unconfirmed source. However, blue whales (*Balaenoptera musculus*) were sighted in the area later in the day.



reduction of approximately 32 is realized. This data bandwidth reduction, as compared to reporting the full amplitude FFT, makes the real-time distribution of spectrogram data from hundreds of hydrophones possible. The detection reports are also stored to disk by M3R's central archiving utility. Archive files form a permanent, electronic record of M3R exercises and are used in postprocessing analysis.

Data Association and Localization

M3R uses hyperbolic multilateration to localize cetaceans. Multilateration requires reception of a given, specific signal by multiple sensors whose positions are precisely known (1). The time difference of arrival (TDOA) of the signal as received at the different sensors is then used to solve a linearized system of equations derived from (1) to determine position (Ludecke, 1998; Vincent, 2001).

$$(x_w - x_i)^2 + (y_w - y_i)^2 + (z_w - z_i)^2 = c_i^2 (t_i - t_e)^2 \text{ for } i = 1 \text{ to } N. \quad (1)$$

where

x_w, y_w, z_w = unknown coordinates of the whale;
x_i, y_i, z_i = known coordinates of the i^{th} hydrophone;
t_i = known time of arrival at the i^{th} hydrophone;
t_e = unknown time of emission of the vocalization;
c_i = computed effective sound speed for travel from the whale to the i^{th} hydrophone; and
N = number of hydrophones

The lack of *a priori* knowledge of the call signals received by the hydrophones presents a major challenge in using multilateration to localize whales. Across any time window, there can be multiple calls from multiple species and/or multiple individuals from the same species. M3R must automatically recognize the arrival of (multiple) signals of interest on sets of several hydrophones. It also must properly associate the signals received across hydrophones to determine TDOA. Localization requires association of detections on a minimum of three noncolinear hydrophones to calculate a 2-D position (a fixed depth is assumed) and on a minimum of four noncolinear hydrophones in order to compute a 3-D position. Using the minimum number of receiving hydrophones can, reportedly, result ambiguities under certain sensor field geometries (Spiesberger, 2001). M3R uses relatively dense sensor fields where detected animals can be assumed to be within a favorable geometry to the set of sensors on which they were detected, or else they would have been detected on a different group of sensors. Based on this proximity assumption, solutions outside the surrounding hydrophone array can be discounted.

M3R performs automated detection association across sets of five to seven hydrophones, which are arranged in either an X or hexagonal pattern. The hydrophone in the center of the pattern is designated as the master hydrophone. All TDOAs are calculated relative to the time of detection on the master phone. The gist of the association problem is to automatically recognize the same, unique pattern of vocalizations, as it was received by the master hydrophone, on the other hydrophones in the array. For the purposes of detection association, M3R divides vocalizations into two broad categories, clicks and whistles, which are associated in different ways.

Click detection and association are performed first. A click is detected if the number of bins above threshold in the full-band FFT detection report exceeds the click threshold, $Thr_{\text{clk}} = 10$. This threshold was determined empirically and results in a measured click false alarm rate of <1 per 100 s (Ward et al., 2008). A probability of click detection, $P_{\text{det}} = 50\%$, is achieved at a received signal-to-noise ratio of approximately 10 dB (Ward et al., 2011). For Blainville's beaked whales at AUTEC, who have an estimated on-axis source level of in excess of 200-dB re µPa (Ward et al., 2011), this translates into a maximum detection range of approximately 6,500 m in low sea state condition (Ward et al., 2011). This detection range agrees with antidotal observation from M3R exercises. The detection ranges for other species are greater or less than this range depending on their respective source levels. Sperm whale clicks, which can have a source level of up to 236-dB re µPa (Møhl et al., 2003), are regularly detected at ranges of 10+ km at AUTEC. Observations from M3R operators at SCORE and PMRF indicate a detection range for dolphin clicks of approximately 4–5 km.

Although the type of call signals that might be received at any time cannot be predicted, the same pattern of calls over time will be received by any hydrophones within the hearing radius. Echolocation foraging clicks are produced by several species in nearly periodic click trains (Johnson et al., 2004; Watwood et al., 2006). A number of species also generate homing pulses (e.g., sperm whale creaks or beaked whale buzzes), which are periods of very rapid clicking produced prior to prey capture attempts (Madsen et al., 2005). Some animals (e.g., dolphins) also produce

social clicks (Lammers et al., 2003). M3R uses the fact that odontocete clicks are produced in patterns to automatically associate click detections among hydrophones neighboring the master. Figure 5a shows a synthetic example of the reception of two interleaved click trains as they are received by the master hydrophone. However, in general, the number of animals is unknown. Any given receiver simply receives a series of clicks (Figure 5b), and it is unclear if or how these clicks relate to the click patterns received by the master hydrophone.

M3R also uses that fact that different species produce clicks in different frequency bands. Analysis of visually verified recordings of animals from AUTEC shows that the clicks produced by different species often have their signal energy concentrated in different bands. For example, most of the energy in a sperm whale click is in the band of 2–18 kHz. Blainville's beaked whale clicks have most of their energy above 24 kHz, while multiple species of dolphins appear to click in the region above approximately 12 kHz. Each click received on the master hydrophone was emitted by a particular animal and is part of only one click train. After a click is detected, it is assigned to one of five frequency regions or click types—LF clicks (<1.5 kHz), sperm whale clicks (1.5–18 kHz), dolphin clicks (12–48 kHz), beaked whale clicks (24–48 kHz), or HF clicks (>45 kHz)—based on the number of bins above threshold in each frequency region. This frequency segmentation provides only a very coarse classification capability. The click types are broad in efforts to maintain generality. More sophisticated classifier processes can be attached to an M3R cluster to provide separate, robust, species level classification (Jarvis, 2012; Jarvis et al., 2008; Martin et al., 2013).

The sequence of click detections within the same frequency band as the first click received on the master channel is used as a template or a click map. The click detection patterns in the same band from the surrounding hydrophones are compared to the click map at different time delays, and the number of matching clicks is tallied. The click map time window always starts on a click detection and is moved across the scanned signal one click detection at a time. The duration of the click map time window is selectable and is currently set to 10 s. The resultant correlation value at any time delay represents the number of click matches between the master channel template and the scanned channel (Figure 6). The delay corresponding to the maximum correlation value represents the difference in time of arrival between the two sensors for that click train.

The output of the click map correlation process is raw sets of TDOA for each click detection received on the master channel. The raw TDOA data from each hydrophone are then tracked over a 5-min history to estimate the number of separable sources. Only detections whose TDOAs lie on a consistent track of time delay versus time are considered (Figure 7). To determine if a consistent track exists, the algorithm sums the number of detections that lie on along the same TDOA as the TDOA associated with the most recent detection. A fixed amount of drift in

FIGURE 5

(a) A notional example showing the patterns of clicks from two sources as they are received on the master hydrophone A. (b) The same two patterns of clicks are also received by hydrophone B, but it is unclear which clicks are associated with which source.

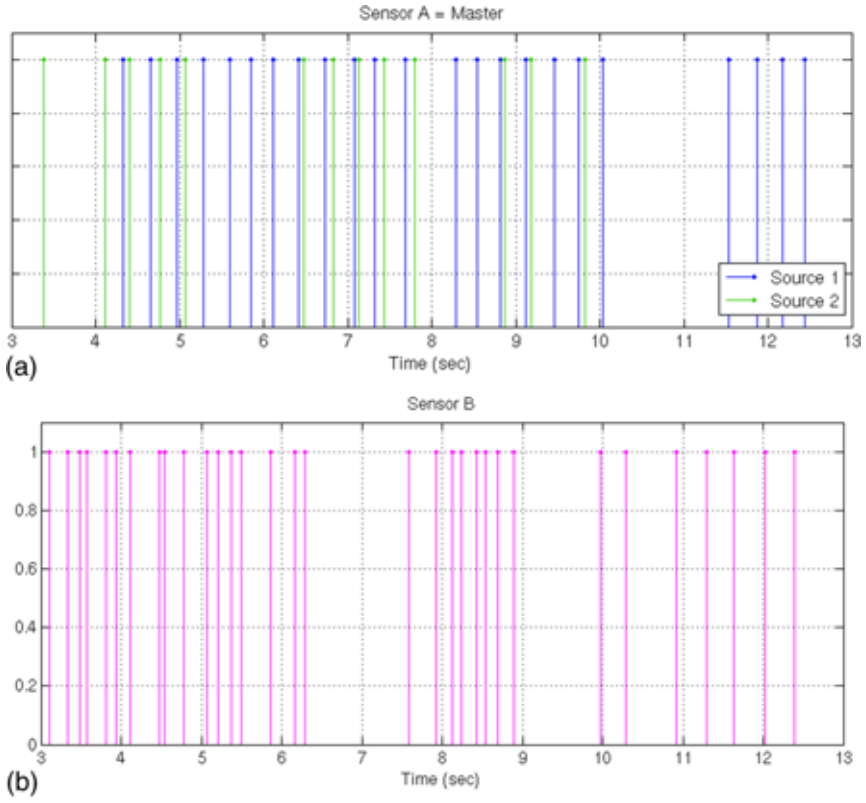
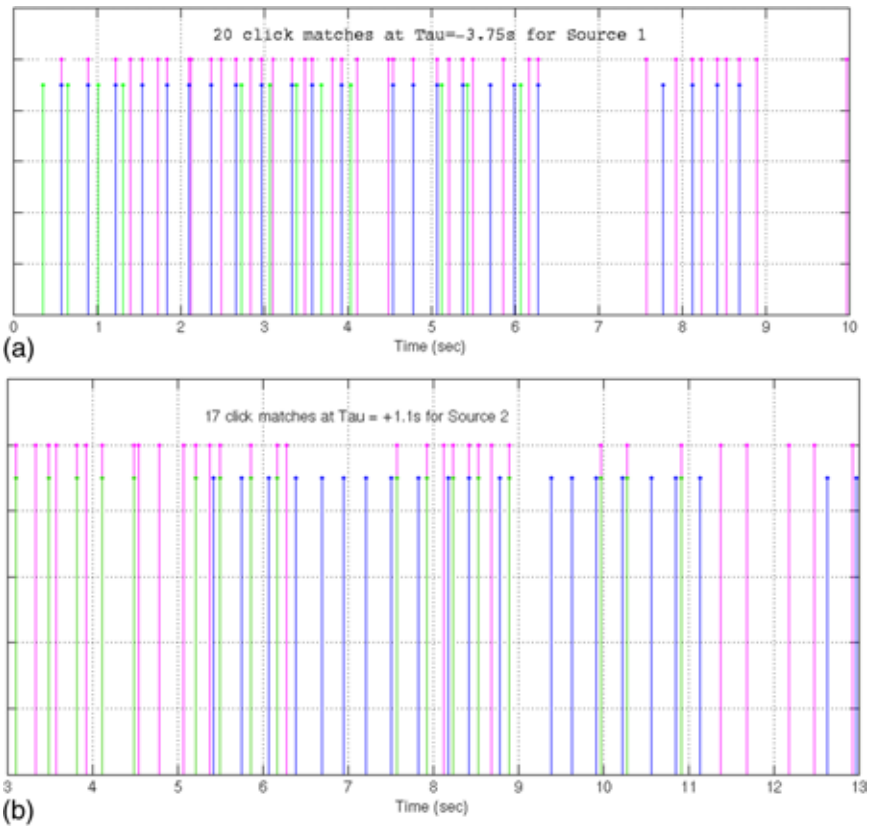


FIGURE 6

(a) Using the clicks as they are received on hydrophone A as a click map and correlating against the click arrivals on hydrophone B separate the clicks from two sources. Source 1 (blue) has TDOA between B and A of –3.75 s, and (b) Source 2 (green) has TDOA between B and A of 1.1 s.



time delay versus time is allowed. The number of detections along the TDOA path is compared to the total number of detections (within that frequency band) in the 5-min history. The TDOA is deemed valid if the ratio of the number of detections, which lie along the TDOA to the total number of detections within the history, exceeds an empirically set threshold. Validated sets of TDOA, which include three or more hydrophones, are then sent to the multilateration localization algorithm. For loud species like sperm whales, multipath detections can cause associations at an erroneous TDOA offset from the true direct path detections. These tracks of TDOA from multipaths are visibly less distinct than the corresponding direct path tracks in manual analysis (Figure 7) but can be hard to reject in automated processing even by tracing the TDOAs through a history (Baggenstoss, 2011). When validated TDOA sets include delays from multipath receptions, the resulting localization (if a solution is found) will be offset from the true animal position. If loud and consistent multipath signals are received from an animal, the automated association algorithm can mistake the multipaths for a second animal and generate false TDOA and localizations for it. For most odontocete clicks, other than sperm whale clicks, such multipath reception on the ranges' deep-water (~1600+ m) hydrophones is typically not an issue.

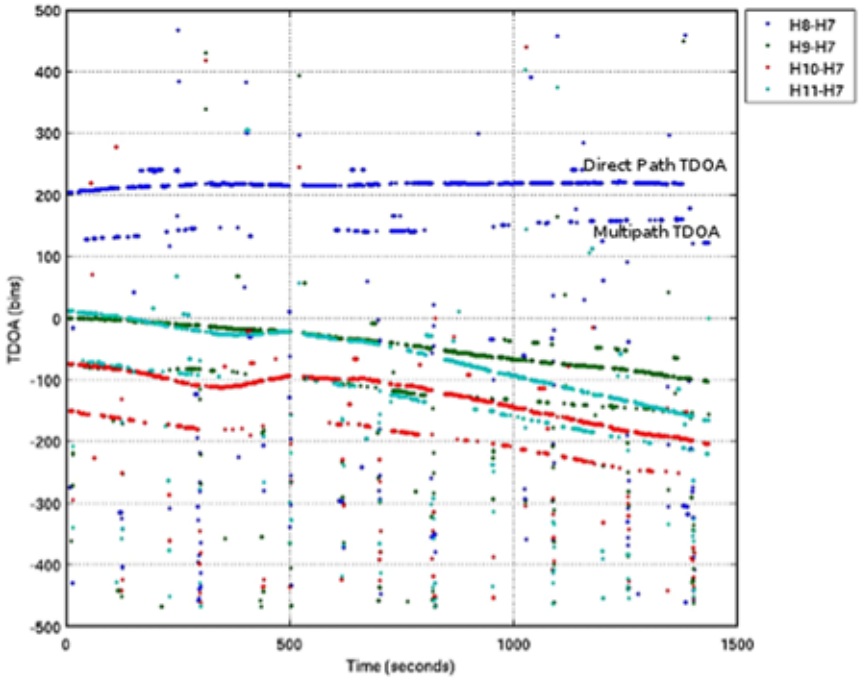
Narrow-band whistle vocalizations are not necessarily emitted in repetitive patterns. An individual can emit a single short whistle or groups of sweeps that last several seconds or both. However, the time-frequency characteristics of the calls in whatever sequence they may occur remain the same on all receiving hydrophones. To determine the TDOA of whistle signals among the hydrophones, the detection spectrograms $Q(f,n)$ of the neighboring hydrophones are cross-correlated against a master channel M. The cross correlation $C_j(n,\tau)$ between the j^{th} channel and the master channel is calculated over a time window of 10 s:

$$C_j(n,\tau) = \Sigma_f \Sigma_{l=n}^{l=n+w} Q_M(f,l) Q_j(f,l-\tau)$$

where w is the number of detection spectra in the 10-s time window and τ is the time delay. The time window is then advanced by 5 s, and $C_j(n,\tau)$ is updated. The time delay associated with the peak of the correlation functions indicates the TDOA for a signal relative to the master hydrophone. If whistles from multiple animals are present within a cross-correlation time window, multiple correlation peaks will be evident (Figure 8). A time-varying threshold based on the mean correlation value across all time delays is applied. Only the time delays associated with peak values above threshold are preserved. Note that, if both clicks and whistles are present within the same 10-s analysis window, the energy from the clicks will dominate the detection spectra. Correlation peaks due to whistle signals will be obscured. Thus, for practical purposes, when a click is detected (i.e., when the number of bins above threshold in the detection spectrum exceeds Thr_{clk}), all

FIGURE 7

Preliminary output of the M3R click association algorithm for a single sperm whale showing the estimated TDOAs between the click as received on master hydrophone 7 and four additional hydrophones. Only TDOAs that are part of a track versus time are passed to the localization algorithm. Spurious TDOAs are ignored. Also, notice the less prominent TDOA tracks, which parallel the direct paths caused by multipath arrivals.



bins of that detection spectrum are set to 0 prior to calculating $C_j(n,\tau)$.

While cross-correlation of detection spectra indicates times of signal arrival and the presence of multiple whales, it does not associate the time delays of the correlation peaks with an individual animal across hydrophones. However, as mentioned earlier, the sequence of whistles versus time from an individual is the same across all receiving hydrophones. Figure 9 shows the time differences of arrival of whistle detections relative to a master hydrophone for five hydrophones. Notice that there are two distinct patterns of detections versus time along specific time delays. Matching these patterns along time delays over a 5-min history associates the TDOAs among the hydrophones with an individual whale. Associated sets of TDOA can then be sent to the multilateration algorithm.

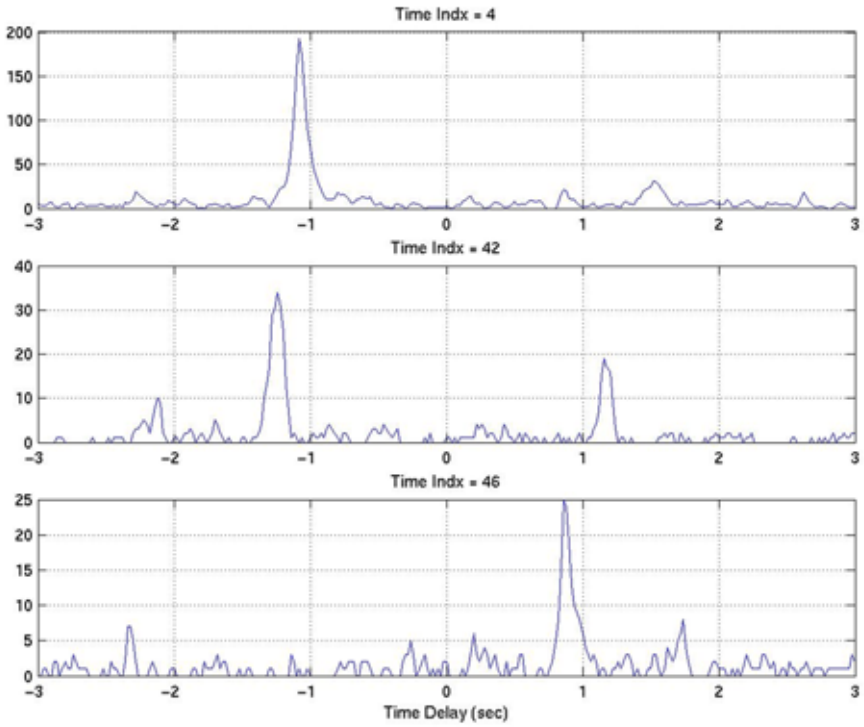
Each TDOA association and validation process within M3R processes data from a single array of five to seven hydrophones. Multiple versions of the associator process are executed in parallel for overlapping arrays spanning the entire range. In an M3R cluster, two to five processing nodes are dedicated to detection association and TDOA validation, and a separate node is dedicated to localization. All raw and validated TDOA sets as well as any resultant localizations are archived to the M3R central archive.

Real-Time Display

M3R provides a number of outputs that can be monitored by operators in real time. The MMAMMAL display (Figure 10) gives an overview of acoustic activity across the entire range. The number of FFT bins above threshold is tallied for each hydrophone and is used as a proxy for vocal activity. The amount of vocal activity is color-coded into the hydrophone number on the display (i.e., the red hydrophone has the most activity; green hydrophones have less activity). The display is interactive, and an operator can open a scrolling "strip chart" of the detection spectrograms from any hydrophone by clicking on the hydrophone number (Figure 11). Whale position data can be displayed through two interactive mapping programs: GoogleEarth (http://www.google.com/earth/index.html) and WorldWind (http://worldwind.arc.nasa.gov/java). The positions automatically calculated through multilaterization are served out (to the public network) through a web server in a Google keyhole markup language (.kml) file. In addition, the positions are archived to the M3R central archive. Users can open a network connection to this .kml file from Google Earth (https://developers.google.com/earth/articles/earthapikml). The result is a display of hydrophone and whale positions overlaying Google Earth images and navigation (Figure 12). Worldwind is an open-source global mapping tool originally developed by NASA. The main benefit of Worldwind (http://goworldwind.org) is that, unlike GoogleEarth, any of its images can be stored on the local machine and it does not require an Internet connection to retrieve images. Also, the Worldwind display (Figure 13) reads M3R position messages multicast by the localization process directly and does not require the creation of separate .kml files.

FIGURE 8

(Top) Results of cross-correlation indicating the TDOA of the signal from one whale. (Middle) Cross-correlation results indicating the TDOAs of signals from two whales. (Bottom) A short time later, only a single correlation peak is visible.



Data Archiving and Postprocessing Analysis Tools

All the detection, association, and localization processes within M3R communicate their results by multicasting report messages on the processed data network (Figure 1). All these messages are stored by a central archiving utility program. This means that an analyst can play back an archive file on a single computer and retrieve all M3R detection spectra, association results, and localizations for the entire range area. Furthermore, since 2006, these archives have been recorded almost continually at AUTEC and SCORE. The archived data can be displayed using the tools discussed above in the Real-Time Display section or postprocessed offline. Recently, M3R data have been used to generate density estimates for Blainville's beaked whale (*Mesoplodon densirostris*) at AUTEC (Marques et al., 2009;

FIGURE 9

TDOA data resulting from spectrogram cross-correlation of five hydrophones against a master hydrophone. Two individual whales show two distinct patterns of whistle detections vs. time along specific TDOAs. The colors of the icons correspond to the different hydrophones in the array and the icon shape (dots and triangles) indicates the two whales. Black Xs are the arrivals that showed no detection pattern across hydrophones.

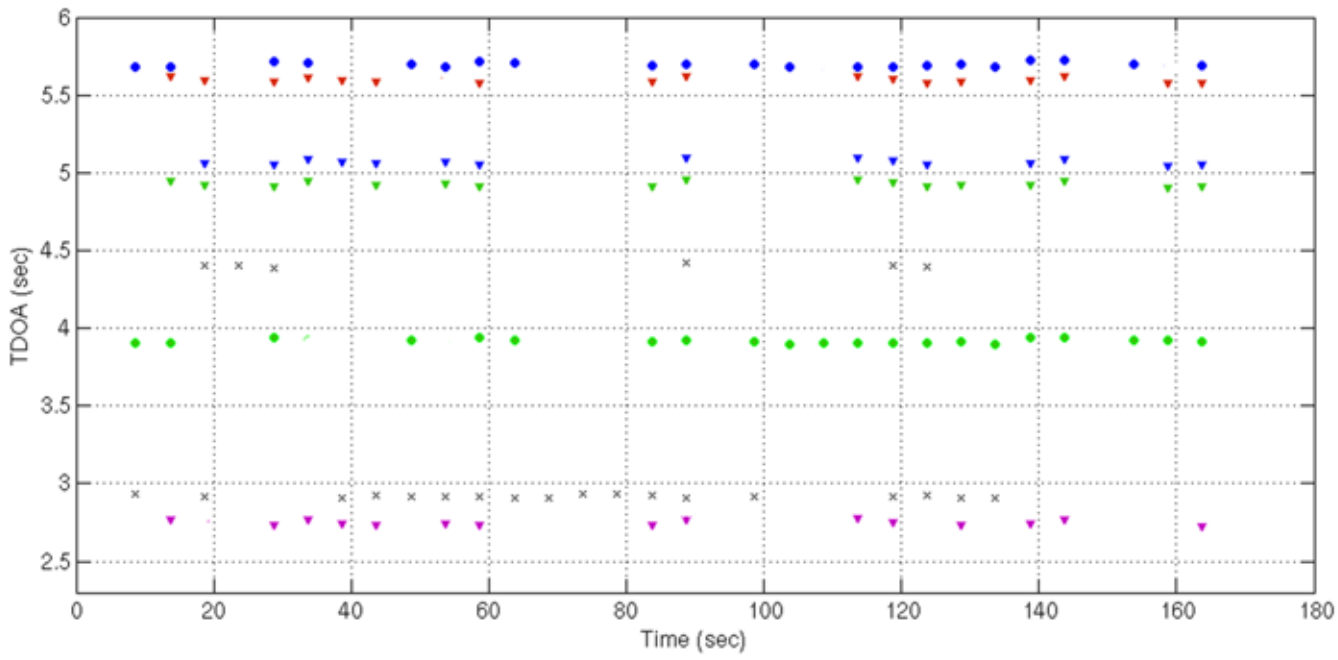


FIGURE 10

The MMAMMAL display provides an overview of acoustic activity over the entire range. The color of the hydrophone number indicates the number of FFT bins above threshold relative to all hydrophones.

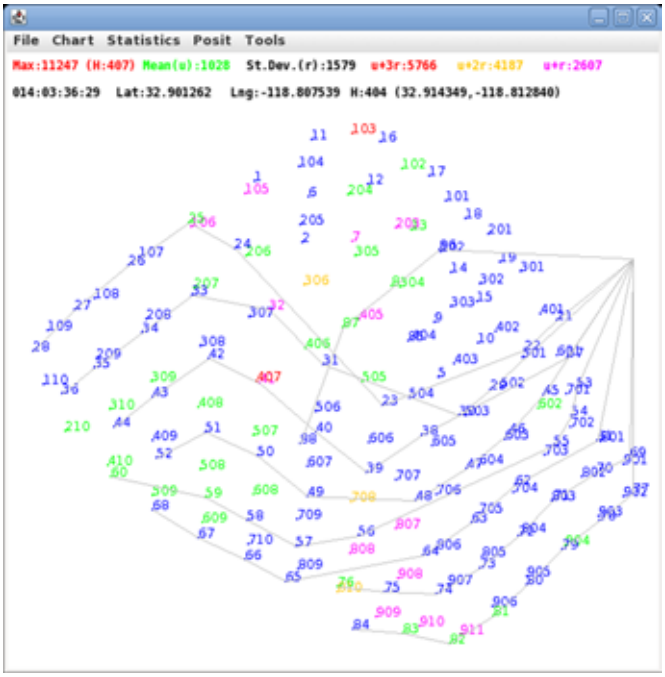
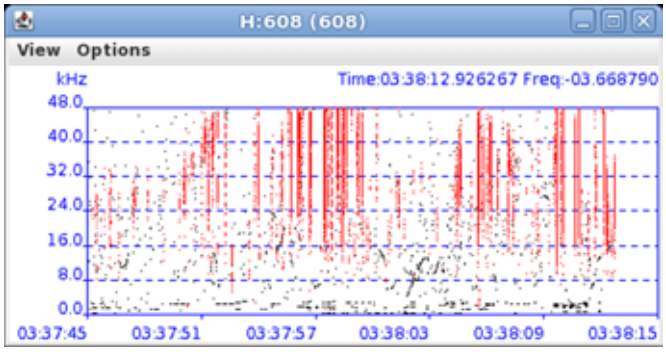


FIGURE 11

Scrolling strip chart displays of detection spectra versus time can be opened by clicking on a hydrophone number in MMAMMAL. This strip chart shows patterns of clicks and whistles associated with multiple common dolphin (*Delphinus delphis*).



Moretti et al., 2010). As discussed in the Data Association and Localization section, M3R click detections undergo frequency segmentation and are assigned to one of five click types. Postprocessing tools group clicks of each type into click trains based on heuristic rules such as interclick interval and dive behavior. For example, Blainville's beaked whales are known to have an interclick interval of approximately 0.25–0.4 s and generally emit clicks at depths of >400 m (Tyack & Johnson, 2006). Counting of click trains and other vocal events established through postprocessing of M3R data have been used as cues for species presence in density estimation (Marques et al., 2010; Moretti et al., 2010) and measuring reaction to anthropogenic sound (McCarthy et al., 2011; Tyack et al., 2011).

M3R also has the ability to record the raw acoustic data from all range hydrophones using a hard disk recorder. At AUTEC, the recording system consists of a bank of eight Alesis HD24 digital hard disk recorders recording at 96 kHz. For SCORE and PMRF, a custom hard disk recorder was implemented, which runs on a single node in the M3R cluster. This recorder stores all raw data packets from the raw data network to standard 2-TB hard disk drives.

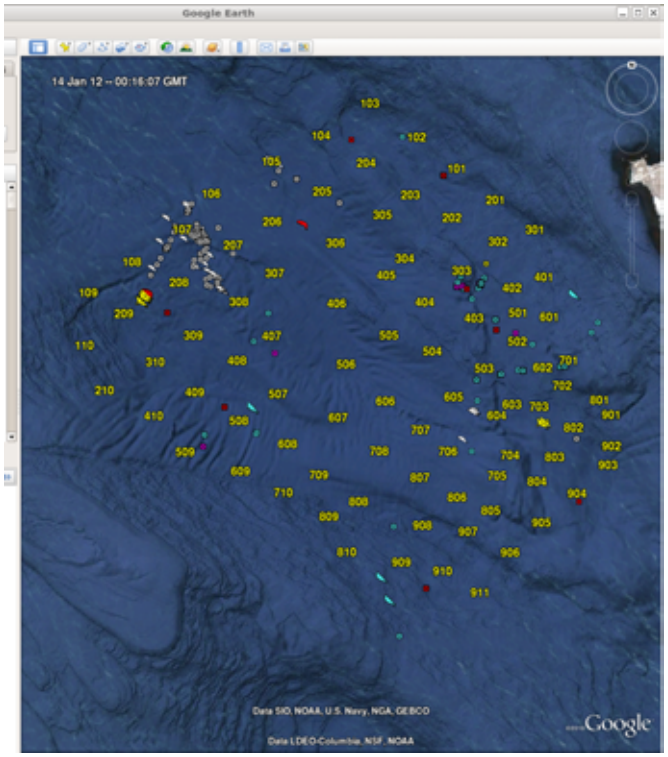
Results

Species Verification

The major challenge in the development of M3R was associating whale vocalizations as they are seen in the detection spectrograms to the correct species. Beginning in 2002, a series of species validation test have been conducted on all three U.S. Navy ranges. Visual observers, usually in small boats, are directed to vocalizing whales based on localizations generated automatically by M3R. The observers relay species information back to M3R, recordings of the hydrophones nearest to the observation are made, and screen shots of the detection spectrograms are taken. Some of these species verification tests have involved tagging individual animals with digital tags (DTAGs) (Johnson et al., 2004) or with satellite tags (Wildlife Computers, 2013). DTAGs are audio recording tags and provided invaluable data for identifying vocalizations. Through these exercises, multiple species of marine mammals have

FIGURE 12

Google Earth display showing M3R whale localizations. Whale-shaped icons are used for the most recent positions (less than 5-min old), while dots are used for positions from 5- to 30- min old. A large number of posits in the same or nearby location indicate higher confidence in the posit.



been visually verified as being present at each of the three ranges (Baird et al., 2013; Claridge & Dunn, 2013; Falcone et al., 2009). Table 1 lists the species verified as present at AUTEC, at SCORE, and at PMRF.

Through these exercises, we have been able to relate detection spectrogram representations to species (Figure 14). Although M3R FFT processing is only capable of coarsely separating animals into six whale types (five click types and whistles), M3R operators can now reliably direct tagging or survey boats to specific species based on a combination of M3R-generated localizations and the MMAMMAL spectrogram display. This M3R-directed approach has been highly successful in tagging specific species in specific study areas such as recent tagging of multiple false killer whales at PMRF (Baird et al., 2013), tagging of multiple Cuvier's beaked whale (*Ziphius cavirostris*) at SCORE (Schorr et al., 2012), and tagging of two Blainville's beaked whale ahead of a multiship sonar exercise at AUTEC.

Localization Accuracy

The M3R monitoring, detection, and localization algorithms were originally designed for and tested at the AUTEC range. AUTEC has 93 long baseline, wide-bandwidth hydrophones mounted in the Tongue of the Ocean, a deep water canyon adjacent to Andros Island, Bahamas. The M3R tool set was later adapted to other ranges (SCORE and PMRF), which have a similar layout of hydrophones spaced on long baselines.

In order to verify the localization algorithms, an experiment was conducted in July 2002 with a calibrated source. Synthetic whale vocalizations were played through an omnidirectional transducer deployed from the stern of a GPS-equipped ship as it drifted beam to the sea through the AUTEC range. The GPS antenna was mounted over the ship's bridge, and the ship's length from bow to stern is 61 m. Both the click and whistle calls generated were localized to within 52 m of the GPS (Figure 15). When corrected for the distance between the GPS antenna and the transducer assembly, 92% of the localizations from the M3R system were found to be within ±15 m of the GPS reading. This positional accuracy is consistent with the uncertainty due to the time resolution of the FFT detection algorithm as the full-band FFT has a time resolution of 10.67 ms and the average speed of sound at AUTEC is approximately 1,504 m/s.

For the purposes of survey or finding animals for tagging, getting the observers to within 100–200 m is usually sufficient. Some odontocetes like beaked whales and sperm whale click almost exclusively during deep foraging dives and are silent during their ascent and while at the surface (Johnson et al., 2004). Thus, M3R localizations for these species are for the animals at depth. During survey or tagging exercises, these M3R positions are relayed to the observer boat via radio then the boats wait at the position for the animal to go quiet and surface. The tag boats frequently remark that, when the animals do surface, they are within a few hundred meters of the position given by M3R. Dolphins and baleen whales often vocalize at or near the surface. When tag boats are directed to localizations for these species, they are

FIGURE 13

Worldwind display showing M3R whale localizations. The Worldwind view also incorporates the FFT detection statistics (total vocal activity) into the color of the hydrophone number similar to the MMAMMAL display.

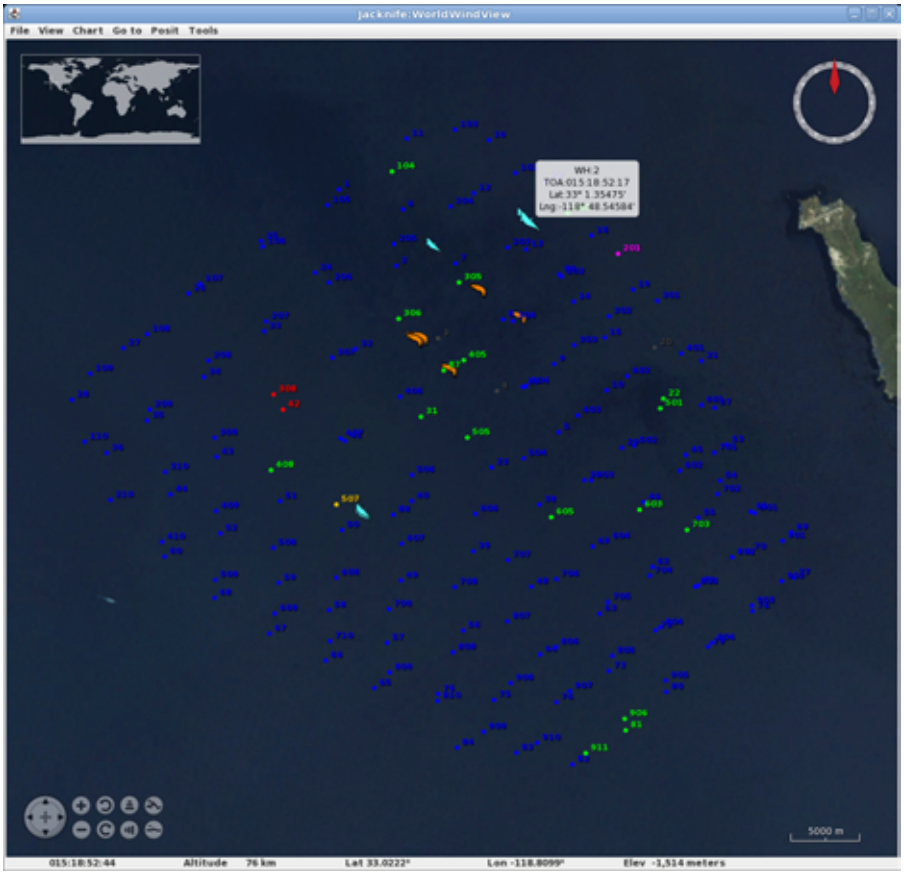


TABLE 1

Species verified as present on each range.

AUTEC	SCORE	PMRF
Sperm whale	Common dolphin	Blainville's beaked whale
Short-finned pilot whale	Risso's dolphin	Cuvier's beaked whale[a]
Rough toothed dolphin	Pacific white sided dolphin	Rough toothed dolphin
Melon-headed whale	Bottlenose dolphin	Bottlenose dolphin
Blainville's beaked whale	Cuvier's beaked whale	Minke whale
Cuvier's beaked whale	Blue whale	Sperm whale
Pan-tropical spotted dolphin	Fin whale	False killer whale
Gervais' beaked whale	Humpback whale	Spinner dolphin
Humpback whale	Killer whale (off-shore)	Humpback whale
Bottlenose dolphin	Baird's beaked whale	Short-finned pilot whale
	Sperm whale	Sperm whale[a]

[a]Acoustic verification only.

likely to see the animals immediately or after a short search.

Monitoring Over Time

The M3R system archives the outputs of all of its FFT detection, data association, and localization processes to a central archive on a USB hard disk. These archives can be played back through any of the M3R analysis and display tools. Furthermore, archives have been recorded nearly continuously at both AUTEC and SCORE since 2006 and at PMRF since 2011. Analysis of M3R archive files has been conducted to show the reaction of Blaineville's and Cuvier's beaked whales to mid-frequency active (MFA) sonar (McCarthy et al., 2011), to measure animal population densities on the ranges (Marques et al., 2010; Martin et al., 2013), and to investigate the population consequences of acoustic disturbance (Moretti, 2011; Moretti et al., 2014). M3R has enabled and supported a number of major scientific exercises including the Behavioral Response Studies in 2007 and 2008 (Boyd, 2009; Tyack et al., 2011) as well as the SOCAL-10 through SOCAL-13 exercises (Southall et al., 2011). The results of these exercises are helping to provide understanding and shape policy regarding the impact of acoustic disturbance on marine mammals.

Conclusions

M3R uses spectrogram-based algorithms for the passive detection and localization of marine mammals using widely spaced, bottom-mounted hydrophones characteristic of U.S. Navy undersea tracking ranges. These algorithms have been implemented and deployed at three of the U.S. Navy's major undersea ranges—AUTEC, SCORE, and PMRF—and

FIGURE 14

Examples of M3R detection spectra (strip chart) displays for several visually verified species: (a) Blainville's beaked whale, (b) a single Sperm whale, (c) multiple common dolphins, and (d) a fin whale.

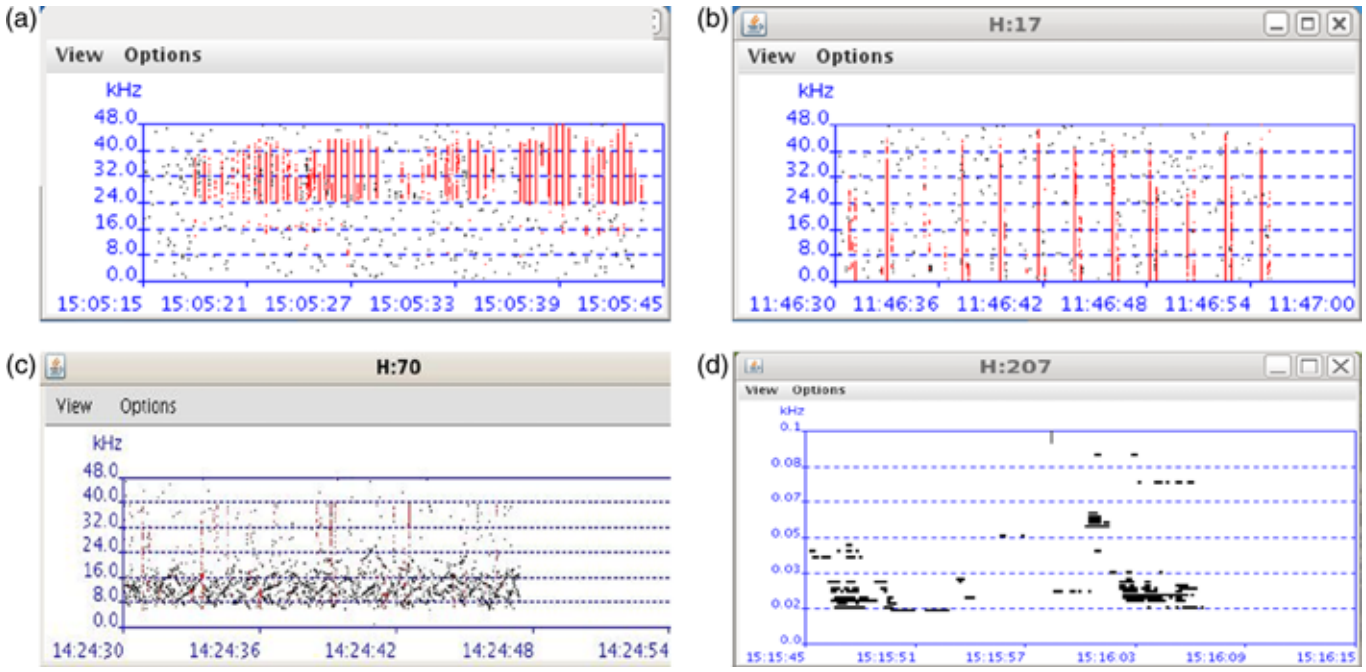
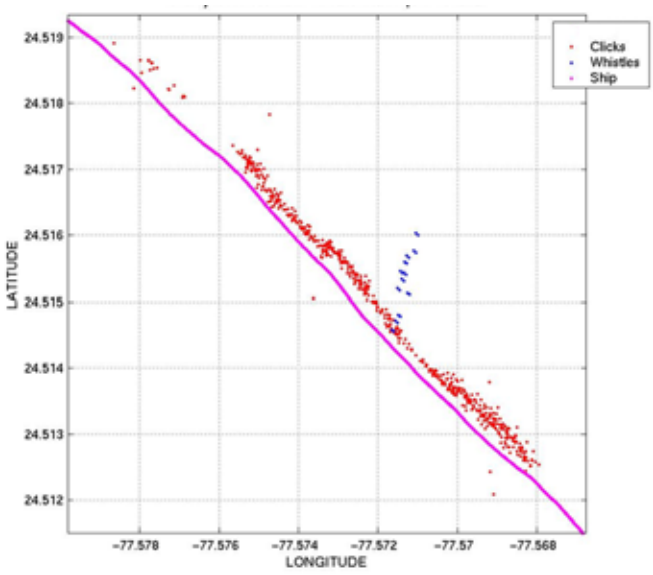


FIGURE 15

Accuracy of the positions generated from M3R for synthetic whale calls. The solid magenta line shows the GPS track of the tow ship, and the dots are the positions generated by M3R. The offset from the GPS track to the M3R posits is approximately 52 m, which corresponds to the approximate distance from the GPS antenna on the ship's bridge to the transducer deployed over the stern.



can monitor over 2,000 km^2 of open ocean in real time. The M3R algorithms have been designed to work in a highly channelized, cluster-based hardware environment, and the software architecture has been developed to be fully network compatible. Signal detection and detection-association algorithms for two primary types of marine mammal calls, whistles and clicks, have been developed. These algorithms are specifically designed to be used with widely spaced sensors. The fully automated detection association and localization algorithms for both clicks and whistles have been demonstrated for calls from several odontocete species including sperm whales (*Physeter macrocephalus*), beaked whales (*Mesoplodon densirostris* and *Ziphius cavirostris*), a number of dolphin species, melon-headed whales (*Peponocephala electra*), false killer whales (*Pseudorca crassidens*), short-finned pilot whales

(*Globicephala macrorhynchus*), and killer whales (*Orca orcinus*). These algorithms require marine mammals to vocalize repetitively with sufficient source levels to be detected on multiple hydrophones. M3R also includes several interactive displays for visualization of animal vocalizations and location.

The M3R system provides tools to collect data previously unavailable for the long-term monitoring of marine mammal bioacoustics within their natural environment. This opportunity has been created with minimal investment in infrastructure by using U.S. Navy ranges as dual-use assets. Research applications of the M3R system include the ability to remotely estimate beaked whale abundance, the assessment of bioacoustic behavioral baselines, and the evaluation of the impact of anthropogenic noise compared to these baselines. The systems also provide a test bed for the development and evaluation of detection, classification, localization, and density estimation algorithms and tools. The data from M3R PAM combined with data from visual observations are being used to derive both movement and population health models for species including beaked whales resident on U.S. Navy ranges where the animals are at risk to repeated exposure to MFA sonar.

Acknowledgments

We would like to acknowledge continuing support from CNO N45 (Dr. Frank Stone) through the Navy's Living Marine Resources Program (Dr. Robert Gisiner) and from the Office of Naval Research (Dr. Michael Weiss). We would also like to acknowledge fruitful ongoing collaborations with Diane Claridge and Charlotte Dunn and the researchers of the Bahamian Marine Mammals Research Organization along with John Durban and Robert Pittman from the National Oceanic and Atmospheric Administration and with the researchers of Cascadia Research Collective including Greg Schorr, Erin Falcone, Robin Baird, Daniel Webster, and John Calambokidis. Also, we thank Peter Tyack and Mark Johnson from University of St. Andrews. Finally, we would like to acknowledge the continued support from AUTEC, SCORE, and PMRF in allowing M3R access to the ranges.

Corresponding Author:

Susan M. Jarvis
Naval Undersea Warfare
Center Division
1176 Howell Street,
Newport, RI 02841
Email: susan.jarvis@navy.mil

References

Au, W.W.L. 1993. The Sonar of Dolphins. New York: Springer-Verlag. pp. 98-177. http://dx.doi.org/10.1007/978-1-4612-4356-4_6.

Baggenstoss, P.M. 2011. Separation of sperm whale click-trains for multipath rejection. J Acoust Soc Am. 129(6):3598. http://dx.doi.org/10.1121/1.3578454.

Baird, R.W., Webster, D.L., Mahaffy, S.D., Schorr, G.S., Aschettino, J.M., & Gorgone, A.M. 2013. Movements and spatial use of odontocetes in the western Main Hawaiian Island: Results of a three year study off O'ahu an Kaua'i. Final report under Grant No. N00244-10-1-0048, Naval Post Graduate School, Monterey, CA, http://www.cascadiaresearch.org/hawaii/Bairdetal_NPS_final_report.pdf.

Baumgartner, M.F., Van Parijs, S.M., Wenzel, F.W., Trembly, C.J., Esch, H.C., & Warde, A.M. 2008. Low frequency vocalizations attributed to sei whales (*Balaenoptera borealis*). J Acoust Soc Am. 124(2):1339-49. http://dx.doi.org/10.1121/1.2945155.

Boyd, I.L. 2009. Behavioral Response Study, ONR Project Report. Available at www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA531250

Claridge, D., & Dunn, C. 2013. Bahamas marine mammal research organization sightings index. http://bahamaswhales.org/sightings/index.html.

Clark, C. 2012. Cornell Lab of Ornithology, "Undersea recording: Pop-ups." http://www.birds.cornell.edu/brp/hardware/pop-ups.

Cox, T.M., Ragen, T.J., Read, A.J., Vos, E., Baird, R.W., Balcomb, K., . . . Benner, L. 2006. Understanding the impacts of anthropogenic sound on beaked whales. J Cetacean Res Manag. 7(3):177-87.

England, G.R. 2001. Joint Interim Report Bahamas Marine Mammal Stranding Event of 15-16 March 2000. www.nmfs.noaa.gov/pr/pdfs/health/stranding_bahamas2000.pdf.

Falcone, E.A., Schorr, G.S., Douglas, A.B., Calambokidis, J., Henderson, E., . . . Moretti, D. 2009. Sighting characteristics and photo-identification of Cuvier's beaked whales (Ziphius cavirostris) near San Clemente Island, California: A key area for beaked whales and the military? Marine Biology. 156:2631-40. http://dx.doi.org/10.1007/s00227-009-1289-8.

Jarvis, S.M. 2012. A novel method for multi-class classification using support vector machines. Ph.D. thesis, University of Massachusetts, Dartmouth.

Jarvis, S.M., DiMarzio, N.A., Morrissey, R.P., & Moretti, D.J. 2008. A novel multi-class support vector machine classifier for automated classification of beaked whales and other small odontocetes. Canadian Acoustics. 36(1):129.

Johnson, M., Madsen, P.T., Zimmer, W.M.Z., Aguilar de Soto, N., & Tyack, P.L. 2004. Beaked whales echolocate on prey. P Roy Soc. 271(6):S383-6.

Klink, H., Mellinger, D.K., Klink, K., Bogue, N.M., Luby, J.C., Jump, W.A., . . . Baird, R.W. 2012. Near real-time Acoustic monitoring of beaked whales and other cetaceans using a

seaglider™. PLoS One. 7(5):e36128. http://dx.doi.org/10.1371/journal.pone.0036128.

Lammers, M.O., Au, W.W.L., & Herzing, D.L. 2003. The broadband acoustic signaling behavior of spinner and spotted dolphins. J Acoust Soc Am. 114(3):1629-39. http://dx.doi.org/10.1121/1.1596173.

Lammers, M.O., Brainard, R.E., Au, W.W., Mooney, T.A., & Wong, K.B. 2008. An ecological acoustic recorder for long term monitoring of biological and anthropogenic sounds on coral reefs and other marine habitats. J Acoust Soc Am. 123(3):1720-8. http://dx.doi.org/10.1121/1.2836780.

Ludecke, F.G. 1998. Tracking algorithms for the naval tracking and display system (NTADS) software. NUWC-NPT Technical Memorandum 9800035, Naval Undersea Warfare Center, Newport, RI.

Madsen, P.T., Johnson, M., Aguilar de Soto, N., Zimmer, W.M.X., & Tyack, P. 2005. Biosonar performance of foraging beaked whales (*Mesoplodon densirostris*). J Exp Biol. 208:181-94. http://dx.doi.org/10.1242/jeb.01327.

Marques, T.A., Thomas, L., Martin, S.W., Mellinger, D.K., Jarvis, S., Morrissey, R.P., . . . DiMarzio, N. 2010. Spatially explicit capture recapture methods to estimate minke whale density from data collected at bottom mounted hydrophones. J Ornithol (2012). 152(Supl 2):S445-455.

Marques, T.A., Thomas, L., Ward, J., DiMarzio, N., & Tyack, P.L. 2009. Estimating cetacean population density using fixed passive acoustic sensors. J Acoust Soc Am. 125(4):1982-94. http://dx.doi.org/10.1121/1.3089590.

Martin, S.W., Thomas, L., Marques, T.A., Morrissey, R.P., Jarvis, S., Dimarzio, A., & Mellinger, D.K. 2013. Estimating minke whale (*Balaenoptera acutorostrata*) boing sound density using passive acoustic sensors. Mar Mammal Sci. 29(1):142-58. http://dx.doi.org/10.1111/j.1748-7692.2011.00561.x.

McCarthy, E., Moretti, D.J., Thomas, L., DiMarzio, N., Morrissey, R., Jarvis, S., . . . Dilley, D. 2011. Changes in spatial and temporal distribution and vocal behavior of Blainville's beaked whales (*Mesoplodon densirostris*) during multi-ship exercises with mid-frequency sonar. Mar Mammal Sci. 27(3): E206-26. http://dx.doi.org/10.1111/j.1748-7692.2010.00457.x.

Møhl, B., Wahlberg, M., Madsen, P.T., Heerfordt, A., & Lund, A. 2003. The monopulsed nature of sperm whale clicks. J Acoust Soc Am. 114:1143. http://dx.doi.org/10.1121/1.1586258.

Moretti, D.J. 2011. Measuring the population health of sonar exposed beaked whales, an energetics approach. http://mmc.gov/pdf/PCADSymposium_moretti.pdf.

Moretti, D., Thomas, L., Marques, T., Harwood, J., Dilley, A., Neales, B., . . . Morrissey, R. 2014. A risk function for behavioral disruption of Blainville's beaked whales (*Mesoplodon densirostris*) from mid-frequency active sonar. PLoS One.

Moretti, D.J., Marques, T.A., Thomas, L., DiMarzio, A., Dilley, A., Morrissey, R., . . . Jarvis, S.M. 2010. A dive counting density estimation method for Blainville's beaked whale (*Mesoplodon densirostris*) using a bottom-mounted hydrophone field as applied to mid-frequency active (MFA) sonar operation. Appl Acoust. 71:1036-42. http://dx.doi.org/10.1016/j.apacoust.2010.04.011.

Morrissey, R.P., Ward, J., DiMarzio, N., Jarvis, S., & Moretti, D.J. 2006. Passive acoustic detection and localization of sperm whales (*Physeter macrocephalus*) in the Tongue of the Ocean. Appl Acoust. 67(11-2):1091-105. http://dx.doi.org/10.1016/j.apacoust.2006.05.014.

Pavan, G., Fosti, C., Francia, C., & Zardin, F. 2010. Acoustic surveys in the Alboran Sea to map the presence and distribution of Cuvier's beaked whales. J Acoust Soc Am. 128(4):2483. http://dx.doi.org/10.1121/1.3508910.

Perkins, P. 1987. Bioacoustics at AUTEC, Bahamas: A survey and guide, technical memorandum no. 87-2018. Newport, RI: Naval Underwater Systems Center.

Rankin, S., & Barlow, J. 2007. Vocalizations of the sei whale *Balaenoptera borealis* off the Hawaiian Islands. Bioacoustics. 16:137-45. http://dx.doi.org/10.1080/09524622.2007.9753572.

Schorr, G., Daniel Webster, D., & Baird, R. 2012. Using satellite tags to study movements and population structure of Hawaiian odontocetes. http://www.cascadiaresearch.org/hawaii/satellite.htm.

Southall, B., Brau, R., Gullard, F.M.D., Heard, A.D., Baird, R.W., Wilkins, S.M., & Rowles, T.K. 2005. July 2004 mass stranding of melon-headed whales in Hawai'i. http://www.nmfs.noaa.gov/pr/pdfs/health/stranding_melonheadedwhales_final_report.pdf.

Southall, B.L., Calambokidis, J., Tyack, P., Moretti, D., Hildebrand, J., Kyburg, C., . . . Barlow, J. 2011. Project report: Biological and Behavioral Response Studies of Marine Mammals in Southern California, 2010 (SOCAL-10). Available at www.SOCAL-BRS.org/SOCAL10

Spiesberger, J.L. 2001. Hyperbolic location errors due to insufficient number of receivers. J Acoust Soc Am. 109(6):3076-9.

Thompson, P.O., Cummings, W.C., & Ha, S.J. 1986. Sound, source levels, and associated behavior of humpback whales, Southeast Alaska. J Acoust Soc Am. 80(3):735-40. http://dx.doi.org/10.1121/1.393947.

Tyack, P.L., & Johnson, M. 2006. Extreme diving of beaked whales. J Exp Biol. 209: 4238-53. http://dx.doi.org/10.1242/jeb.02505.

Tyack, P.L., Zimmer, W.M.X., Moretti, D.J., Southall, B.L., Claridge, D.E., Durban, J.W., . . . Boyd, I.L. 2011. Beaked whales respond to simulated and actual navy sonar. PLoS One. 6(3):e17009. http://dx.doi.org/10.1371/journal.pone.0017009.

Vincent, H.T. 2001. Models, algorithms and measurements for underwater acoustic positioning. Ph.D. Thesis, University of Rhode Island.

Ward, J., Jarvis, S., Moretti, D., Morrissey, R., & DiMarzio, N. 2011. Beaked whale (*Mesoplodon densirostris*) passive acoustic detection

in increasing ambient noise. J Acoust Soc Am. 129(2):662-9. http://dx.doi.org/10.1121/1.3531844.

Ward, J., Morrissey, R., Moretti, D., DiMarzio, N., Jarvis, S., Moretti, D., . . . White, C. 2008. Passive acoustic detection and localization of mesoplodon densirostris (Blainville's beaked whale) vocalizations using distributed bottom-mounted hydrophones in conjunction with a digital tag (DTag) recording. Can Acoust. 36(1):60-6.

Ward, J.A. 2002. Sperm whale bioacoustic characterization in the Tongue of the Ocean, Bahamas, Naval Undersea Warfare Center Technical Report, NUWC-TR 11398.

Ward, J.A., Fitzpatrick, K.C., DiMarzio, N.A., Moretti, D.J., & Morrissey, R.P. 2000. New algorithms for open ocean marine mammal monitoring. OCEANS 2000 MTS/IEEE Conference and Exhibition. 3:1749-52.

Watkins, W.A. n.d. MBL WHOI, MC-78. William A. Watkins collection of marine mammal sound recordings. http://dla.whoi.edu/manuscripts/node/198116.

Watwood, S.L., Miller, P.J., Johnson, M., & Tyack, P.L. 2006. Deep-diving foraging behavior of sperm whales (*Physeter macrocephalus*). J Anim Ecol. 75(3):814-25. http://dx.doi.org/10.1111/j.1365-2656.2006.01101.x.

Wiggins, S.W., & Hildebrand, J.A. 2007. High-frequency acoustic recording package (HARP) for broad-band, long-term marine mammals monitoring. http://escholarship.org/uc/item/0p6832s1.

Wildlife Computers: Innovative Tags for Innovative Research. 2013. http://www.wildlifecomputers.com/.