A large-aperture low-cost hydrophone array for tracking whales from small boats

B. Miller and S. Dawson

Department of Marine Science, University of Otago, P.O. Box 56, Dunedin 9054, New Zealand

(Received 25 February 2009; revised 13 August 2009; accepted 1 September 2009)

A passive sonar array designed for tracking diving sperm whales in three dimensions from a single small vessel is presented, and the advantages and limitations of operating this array from a 6 m boat are described. The system consists of four free floating buoys, each with a hydrophone, built-in recorder, and global positioning system receiver (GPS), and one vertical stereo hydrophone array deployed from the boat. Array recordings are post-processed onshore to obtain diving profiles of vocalizing sperm whales. Recordings are synchronized using a GPS timing pulse recorded onto each track. Sensitivity analysis based on hyperbolic localization methods is used to obtain probability distributions for the whale's three-dimensional location for vocalizations received by at least four hydrophones. These localizations are compared to those obtained via isodiachronic sequential bound estimation. Results from deployment of the system around a sperm whale in the Kaikoura Canyon in New Zealand are shown. © 2009 Acoustical Society of America. [DOI: 10.1121/1.3238258]

PACS number(s): 43.30.Wi, 43.30.Sf [WWA] Pages: 2248–2256

I. INTRODUCTION

Passive acoustic localization, via arrays of hydrophones, has been used to study marine mammals for over 40 years. Instrumentation used in previous passive acoustic studies of sperm whales has ranged from a single hydrophone, stereo hydrophone arrays, multihydrophone towed and vertical arrays, and hydrophone arrays deployed simultaneously from multiple boats. Recent efforts to lower the entry barriers for scientists interested in passive localization include the system presented by Hayes et al. in 2000. This array of relatively inexpensive passive sonar buoys made mostly from commercially available off-the-shelf components was successfully used to track blue whales over several kilometers. In 2001 Møhl et al. presented a similar unlinked sonar array, but instead of using buoys with single hydrophones, hydrophone arrays were deployed from multiple boats to track diving sperm whales. By deploying a relatively deep vertical array, they were able to reconstruct three-dimensional (3D) sperm whale tracks.

Study design and hardware choice limit which software techniques are appropriate for analysis of the raw data. Beamforming techniques are most appropriately applied for short aperture towed or vertical arrays and usually yield animal locations in one or two dimensions (bearing and range). To obtain 3D animal locations, hyperbolic multilateration is usually applied to a multi-receiver large aperture array. These hyperbolic localization techniques were originally developed for the LORAN system which used radio beacons to localize ships and aircraft, and a key assumption was that the propagation speed of the transmission remained constant between the source and all receivers. The assumption of a constant speed of sound propagation is not always valid for underwater applications, especially in a place like Kaikoura, New Zealand, where several water masses converge and complex physical oceanographic processes are at work.

Many of the more advanced localization algorithms involve acoustic ray-tracing and can make use of acoustic multipath detected in the recordings. Ray-tracing and acoustic multipath localization algorithms require detailed knowledge of the sound velocity profile. Ray-tracing techniques can be computationally intensive especially when the hydrophone positions are not fixed with respect to each other. Additionally, multipath localization algorithms require detailed knowledge of local bathymetry. If computation time is limited, or if detailed bathymetry and sound velocity profiles are not available, then alternative algorithms must be employed instead.

In an effort to improve on hyperbolic localization techniques, Spiesberger introduced a geometric surface called an isodiachron. Isodiachronic localization is a more general form of hyperbolic localization that allows for the effective sound speed between the source and each receiver to differ. By using isodiachrons with sequential Monte Carlo methods, one can estimate not just the location of the sound source, but also the positions of the receivers and effective sound speeds. Spiesberger dubbed this technique sequential bound estimation. Like many Monte Carlo methods, this technique is computationally intensive, but yields greater precision and accuracy than hyperbolic least squares error estimation under some circumstances. Furthermore, sequential bound estimation does not require detailed knowledge of the speed of sound, or the bathymetry, though it can make use of this information if it is available. In the absence of measured sound speed profile (SSP), it allows one to estimate a range of SSPs. Conservative estimates should be accurate (i.e., contain the true value), but may not be as precise as using a measured SSP.

In Sec. II, we present a passive sonar system that can best be described as a hybrid between the approach taken by Møhl et al. and Hayes et al. This system was designed for use from a single small (6 m) boat with the purpose of localizing diving sperm whales which at Kaikoura produce...
loud broadband clicks at a mean click rate of 1.3 clicks/s about 60% of the time (including their time at the surface). Our system consists of four free-floating buoys and one long vertical cables array, all deployed from the same boat. Our system is based on commercially available off-the-shelf hardware, offers ease of deployment and recovery of buoys from a single platform, yet it allows measurements of three dimensional movements of nearby vocalizing whales. The system also allows measurement of bearings to whales that may be a few kilometers away. We describe the approach we have taken in hardware and software, and discuss the advantages and limitations of operating this array from a 6 m boat. Additionally, we show results from deployment of the system around a sperm whale in the Kaikoura Canyon in New Zealand.

II. METHODS

A. Array design

The design was heavily influenced by the limited deck space on a small vessel. The components of the array had to be compact, robust, and quickly deployed. We opted for a modular approach in order to reduce maintenance time (defective components can be quickly replaced) and to ensure that overall success in tracking did not depend too heavily on any one component. The modular approach also leaves open the possibility of adding additional instrumentation as future needs dictate.

Design specifications required sufficient battery and storage capacity to make relatively wideband recordings for several hours at a time. A further requirement common to non-linked arrays is that recordings from different platforms had to be synchronized precisely.

1. Buoys

Similar to the instrument packages of Hayes et al. each of our buoys includes a hydrophone, a recording device, a GPS receiver, a time synchronization device, and a battery pack. Additional instrumentation included an optional fourth order bandpass filter (passband 1–40 kHz), a depth logger attached to each hydrophone, and a VHF locator beacon attached to a small mast on each of the buoys. The recording device used in each buoy was the M-Audio Microtrack 24/96 with a 16 Gbyte compact flash card as the recording medium. Each Microtrack had stereo recording capabilities. We used a sample rate of 96 kHz (16 bit) which gave a maximum record time on 16 Gbyte media of about 6 h. While the Microtrack can record 24 data bits per sample, ambient ocean noise and electrical noise within the device itself effectively rendered this setting superfluous. Recording quality could be lowered to 44.1 kHz to enable up to 10 h of recording, though this was not attempted during this study. As compact flash cards increase in size and decrease in price, recording duration can be increased. The Microtrack recorders were powered using their internal battery which gave up to 8 h of operating time.

All hydrophones were built in-house as described by Barlow et al. A single hydrophone was connected to each buoy with 20–30 m of shielded, harsh-environment ethernet cable. The cable chosen was TMB Proplex CAT5e which has a light Kevlar strength member allowing a maximum working pull of 140 N. Previous experience revealed that recordings made with hydrophones shallower than 20 m, resulted in increased surface noise, as well as distortion from surface echo multipaths. A 2 kg lead weight was attached to the end of the hydrophones to speed deployment, maintain hydrophone depth, and reduce hydrophone drift with respect to the buoy. Each hydrophone was connected to one channel of the recording device via a waterproof connector embedded in the buoy lid, while modulated GPS data were recorded on the other channel.

The GPS used on each buoy was the Garmin GPS-17HVS. This low cost OEM GPS was chosen because of its waterproof housing and its ability to output an accurate timing signal in addition to the raw GPS carrier phase. Via post-processing (with data from a suitable base-station), the raw carrier phase can be used to obtain highly accurate (submeter) position information. The GPS position output was connected to an FSK modulator while the GPS timing signal was connected to an amplitude modulator. Detailed description of the FSK modulation is beyond the scope of this paper, for an overview of FSK modulation with respect to this application consult Mohl et al. An Oceanic Veo 250 personal scuba diving computer was used to record the depth of each hydrophone throughout the duration of the deployment. The depth resolution of the dive computers was 0.3 m. Depth sensors in conjunction with post-processed GPS positions allowed for more accurate estimation of the hydrophone position. This was necessary when ocean currents and/or wind caused the hydrophone to drift so that it was not positioned directly beneath the GPS receiver. Testing at the field site revealed this to be necessary for only the deeper boat-based array.

Within each buoy, the recorder, FSK circuit, band pass filter, and two 12 V gel cell batteries were housed in a custom frame which was made from 80 mm diameter PVC drainage pipe. Slots were cut from the drainage pipe and components secured to the frame via cable ties. The frame was placed within a watertight 100 mm diameter housing also made from PVC drainage pipe. The gel cells were placed at the bottom of the frame with 1 kg of lead ballast to help the buoy maintain a vertical attitude in the water. Closed cell foam was glued around the top of the housing to provide additional buoyancy. Buoy dimensions were 1 m in height and 100 mm in diameter. Deck space on the research vessel was limited so the four buoys were stowed upright on deck in a purpose-built wooden rack. While we chose to package the instruments in a 1 m tall tube, the instruments could have fitted into an enclosure as small
as 0.30 m long, making these buoys especially suitable for operation from small vessels. The extra space inside each buoy can be used for additional instrumentation or extra battery packs.

Hydrophones were secured to the rack next to each buoy, and hydrophone cables were wound onto the rack in a figure eight fashion to prevent tangling. By keeping buoy dimensions small and stowing them upright, deployment remained manageable even with limited deck space. A single manual switch within each buoy activated power to all electronics further facilitating speedy deployment.

2. Boat-based cabled array

The boat-based stereo array consisted of two custom built hydrophones spaced 5 m apart on 105 m of cable. The recording device used on the boat was an Ediorl R-4, four channel digital recorder operating at 96 kHz sample rate with a sample resolution of 16 bits. Hydrophone and GPS data were recorded onto the first three channels, while the fourth channel was used to record dictated commentary about the situation. Commentary included descriptions of animal behaviors, vocalizations, weather information, sea state, estimated whale position, and movements of any other vessels in the area.

Using a deeper array proved crucial for obtaining accurate 3D localizations of the target animal. The short distance between hydrophones of the boat-based array also facilitated tracking individual animals when several were vocalizing at the same time. The boat-based array functioned as a short-aperture vertical array which was used to measure the vertical bearing to vocalizing animals using a custom MATLAB script. For short time scales (tens of seconds), vocalizations with widely different bearings are likely to come from different individuals. Similarly, bearings coming from the same individual would be expected to change gradually over short time scales. These assumptions were used by the classification algorithm to reduce ambiguities during analysis that occurred due to multiple vocalizing animals.

In addition to the recordings made using the stereo hydrophone array, a handheld directional hydrophone and compass were used to estimate the range and bearing to the target animal(s) throughout the deployment. Tracking the whale this way provided feedback necessary to reposition hydrophones that drifted too far from the target whale. Because the depth resolution of our system was provided primarily from the boat-based array, it was especially important that the boat-based array remain close enough to detect the animal continuously throughout the recording. In addition to constant feedback, the range and bearing estimate also provided an independent check on the validity of localizations obtained from the non-linked array.

B. Deployment and recovery

Before deploying the buoys and boat-based array, sperm whales were tracked using a custom-built directional hydrophone. Whales were tracked until they surfaced in order to obtain an identification photograph before deploying the array. The photographed whale (target whale) was typically tracked via directional hydrophone for 20–25 min to ascertain its speed and heading (if any) before deploying the first buoy. Because sperm whales in Kaikoura can travel several kilometers from fluke up, to fluke up, it was important to have an estimate of the target whale’s position and general direction of movement before deploying the buoys.

Buoy were typically deployed in either a triangle or square configuration surrounding the animal(s) of interest, with the boat and stereo hydrophone array (and with a bit of luck the target whale) at the center of the polygon. Typical deployment distances between adjacent buoys were 1.5–2 km. To deploy four buoys typically took 20 min; however, poor weather, other vessels, and navigational hazards such as fishing gear increased deployment time.

To deploy buoys, the boat engine was stopped, the buoy electronics turned on, and the Microtrack set to record. The hydrophone cable was deployed before placing the buoy overboard, and the level meter on the Microtrack was checked to make sure that both the hydrophone and GPS recording chains were functioning and that the recorder gains were set appropriately. The waterproof lid was replaced, and the VHF locator beacon was attached before placing the buoy over the side of the boat. To facilitate recovery, each buoy’s deployment location was marked on the vessel’s navigation GPS.

Some care was required when setting the recording gain for each buoy because this could not be changed after deployment. Over the course of the recording, whale(s) can swim toward some buoys and away from others, causing a large change in received level. The dynamic range of the Microtrack recording unit was not wide enough to accommodate this change in received level without gain adjustment. Had the 24-bit recordings on the Microtrack not been dominated by electrical noise, then changing the bit resolution from 16 to 24 bits would have greatly helped handle this limitation. While hardware-based automatic gain control might have been a possible solution, we did not pursue this. Because the goal of this study was localization of whales, we opted for a higher gain setting in order to detect whales further away and maximize buoy separation. This choice came at the expense of clipping some loud whale vocalizations.

When the target animal was unlikely to be audible on a minimum of four hydrophones, buoys could be repositioned; however, this required an interruption in monitoring with the
boat based array which effectively limited localization accuracy to two dimensions (x and y) during this time. While buoys could be repositioned by a dedicated support vessel, this would effectively double the operating costs of the array and has not been attempted. More often instead of repositioning buoys the recording was terminated and the buoys recovered. After recovery, data were downloaded to a personal computer (PC) for synchronization and analysis.

C. Synchronization

The standard procedure for localizing animals with an acoustic array involves computing time of arrival differences (TOADS) between each pair of hydrophones for each vocalization.\textsuperscript{2,3,14,15,17,18,32–36} When using an unlinked array, all recordings made at each location must be synchronized before TOADS can be computed accurately. Synchronization must address both jitter and clock drift. For the purposes of this article, jitter can be thought of as very short-term changes in the sample rate, while clock drift refers to long-term differences between the device’s clock and the GPS synchronization signal. Both of these errors arise from imperfections in the digital clock used for analog-to-digital conversion in the recording unit.

For our system, jitter, measured during synchronization, was typically on the order of 0.002% for all devices. Clock drift was also measured and was typically between 0.5 and 2 ms/min for all devices. Measurement of the jitter and drift rates is not only necessary for accurate localization but also provides a measure of the temporal fidelity of the audio device. Audio time alignment and jitter/drift correction was performed via a two-stage process. The first step involved coarse alignment, which synchronized the start and end of each recording to within 1 s and assumes constant clock drift and no jitter. The second step (fine scale alignment) provided sample-accurate audio synchronization once every second for the duration of the recorded audio.

The GPS position information stored in the FSK-modulated audio signal included latitude, longitude, as well as UTC date and time of the signal with time resolution of 1 s. For coarse alignment, we extracted this information from the first and last seconds of the recording to compute the GPS start and end seconds for the recording. Subtracting the ending GPS second from the starting second yields the GPS duration, $t_{GPS}$. The average clock drift rate was computed as $(t_{GPS} - t_{recording})/t_{GPS}$, where $t_{recording}$ was the total number of audio samples per channel divided by the nominal sample rate (96 kHz). This coarse alignment does not account for jitter or inaccuracies resulting from a non-constant clock drift rate over the duration of the recording. To investigate these errors and account for them if they are present, we used the Garmin GPS 17 timing signal, which is a 1 Hz pulse wave with a duty cycle of 0.1. The rising edge of this pulse marked the start of the GPS second with a nominal accuracy of ±1 μs (Garmin GPS 16/17 Technical manual). When the 100 ms pulse was active, it reduced the signal amplitude of FSK-modulated GPS data. The instantaneous sample rate of the recording unit was computed by simply counting the number of samples between successive pulse edges, and the instantaneous jitter was computed as the difference between nominal sample rate and instantaneous sample rate.

Custom synchronization software written using Mathworks MATLAB was used to detect the sample number corresponding to the leading edge of the amplitude-modulated timing pulse. This edge detection software began by loading 1 s of audio into memory. This audio was divided into ten consecutive sequences and the rms amplitude of each sequence was computed. Due to the amplitude modulation, the sequences containing the timing pulse had a different rms amplitude than the rest of the signal. The earliest sequence with a different rms amplitude contained the leading edge of the pulse. This sequence was kept, while the others were discarded. For each sample in the remaining sequence, the rms value of the subsequent 20 samples was computed. The difference between sequential rms values was computed and the sample with the largest change in rms amplitude corresponded to the leading edge of the PPS. This process was repeated for the duration of the recorded audio.

Simultaneous to edge detection, the FSK-modulated audio track was played into a custom-made hardware demodulator and the GPS data were recorded via a PC serial port. Hardware FSK demodulation with concurrent software PPS detection allowed for synchronization of multiple buoys at the same time. The audio sample number of the PPS edge, latitude, longitude, UTC time, and raw carrier phase information from each platform was written to a synchronization data file for every second of audio processed (Fig. 2). By using a hardware-based demodulator circuit for each of the five audio channels, 5 channel-hours of modulated GPS position and timing information could be decoded in 1 h. This proved to be significantly faster than our best attempts at implementing software-based demodulation as described by Møhl et al.\textsuperscript{14} and can work in real time provided that there are as many demodulators and serial ports as there are modulated GPS signals.

Hydrophone depth sensors were activated via a water contact switch. Depth data were synchronized with the GPS data using either the audible tone made when the depth sensor was active for a predetermined amount of time or the sound made from the entry of the depth sensor into the water, both of which were audible in the recording for each hydrophone. At the end of preprocessing, the multichannel recordings of the whales’ sounds, location data for each buoy and the boat, and the depth data for each hydrophone were synchronized.

FIG. 2. Data flow for time alignment and demodulation of recorded GPS data.
D. Detection and localization

Recently there have been many different techniques proposed for detection, classification, and localization of sperm whale clicks. To obtain 3D whale positions, we implemented a selection of detection, classification, and localization algorithms using Mathworks MATLAB 7.3, and adapting the methods for use with our system as necessary.

1. Detection

For detection of sperm whale vocalizations, audio recordings from each platform were bandpass filtered between 2 and 20 kHz, a band which contains most of the energy of typical sperm whale vocalizations. Vocalizations were detected from filtered recordings using Page’s test, which is an energy detector. Specifically, we followed the algorithm outlined in Ref. 39, Part II, Sec. I. While there have recently been numerous methods for detection and classification of sperm whale clicks, Page’s test was chosen because its implementation was intuitive, fast, and it has been used successfully in previous studies involving localization of sperm whales. Detection parameters that yielded good agreement with visual inspection of the spectrogram for the first few minutes of audio were selected for use. The detection threshold, \( V_1 \), was set to 16 (24 dB); the end of detection threshold, \( V_0 \), was 1; and the exponential weighting on the noise, \( \alpha \), was set to 0.9 (notation follows Ref. 39). Automating the detection process was necessary for analyzing the large number of recorded sperm whale vocalizations.

2. Bearing localization

TOADs were computed between both hydrophones from the boat-based stereo array using cross-correlation of the waveform of detected clicks. The time lag at the peak in the cross-correlation function was recorded as the TOAD of a direct arrival. Because the distance between these hydrophones was much smaller than the distance to the target whale, these TOADs provide a measure of the angle of arrival of the sound. These angles were plotted as a function of time to yield a bearing-time plot for the recording. Bearing tracks were traced by a human operator and tracks were numbered and assigned to an individual whale via a custom MATLAB interface. Bearings were traced with the following criteria. Bearings that corresponded to an individual whale track must change slowly and continuously over time. This constraint eliminated noise sources from being selected as a whale trace. Any ambiguities in a trace, such as the intersection of multiple traces or gaps longer than 7 min, resulted in the termination of a trace and the start of a new trace at a time after the ambiguities could be resolved. A recording typically contained between 1 and 6 individual bearing traces at any one time.

3. Surface echo detection

Echo detection based on autocorrelation was performed on vocalizations from each bearing trace from the vertical array. For each vocalization, the absolute value of the autocorrelation of the waveform was computed. The largest peak in this autocorrelation function that occurred between 10 and 200 ms after the direct arrival was considered a surface echo so long as the time lag of this peak did not correspond to a direct arrival from another bearing trace. All surface reflections from a particular bearing trace were written to a separate log file. These surface reflections can be thought of as arriving at a virtual hydrophone that mirrors the real hydrophone above the ocean surface. These virtual hydrophones were used as additional receivers and increase both the number of hydrophones in the array and the vertical hydrophone separation, thus increasing the localization performance of the array.

4. Classification (click association)

The inter-click interval from each of the bearing traces was computed and used as input into a custom MATLAB program that implemented the “rhythm analysis” algorithm described by Thode. This algorithm was necessary to associate vocalizations received at each buoy and virtual hydrophone with vocalizations received from an individual whale at the stereo array. When a vocalization was matched at 4 or more real hydrophones, arrival time differences were calculated between all hydrophone pairs by computing the cross-correlation of the audio for each matching detection. The time lags of the peak of the cross-correlation function were stored as TOADs and used in further localization analysis.

5. Localization

Once all TOADs have been computed, these data as well as the hydrophone positions were used as input into a MATLAB program that implemented the hyperbolic localization algorithm described by Spiesberger and Fristrup. To estimate the localization precision, a separate sensitivity analysis was performed.

For the sensitivity analysis, we assigned uncertainty to each of the model inputs and created uniform probability distributions based on the measured data and estimated/measured uncertainty for each of the model inputs. Variance for the horizontal hydrophone position was measured to be ±2 m based on a 48 h comparison of each GPS receiver to a surveyed reference position. Variance for each hydrophone depth sensor was assumed to be ±0.3 m according to the manufacturer’s specifications. Effective sound speeds were allowed to vary across the range of sound speeds computed with equations from Del Grosso using historical monthly temperature, salinity, and depth data for the study area from the World Ocean Atlas. TOAD variance was computed according to Spiesberger and Fristrup equation 41. We then drew 2000 samples from each of these random variables and used each set of samples as input to the localization algorithm to obtain a cloud of points that represents the whale’s position. For the \( x \), \( y \), and \( z \) coordinates of the whales position, probability distributions, \( P_x \), \( P_y \), and \( P_z \), were estimated at each time step from the output of the sensitivity analysis. Estimates of \( P_x \), \( P_y \), and \( P_z \) were calculated from the normalized histograms of each \( x \), \( y \), and \( z \) coordinate of the whale’s position using bin widths of 1 m. The total volume for each
localization cloud was computed as $(\hat{P}_x - \tilde{P}_x)(\hat{P}_y - \tilde{P}_y)(\hat{P}_z - \tilde{P}_z)$, where $\hat{P}$ and $\tilde{P}$ denote the maximum and minimum values from the probability distributions. The total cloud volume is a measure of localization precision. A threshold volume of $1.77 \times 10^6$ m$^3$, which is equal to the volume of a sphere with a diameter of 150 m (approximately 10 whale lengths), was used to exclude localizations with low precision (Fig. 6).

Because hyperbolic localization can yield incorrect results in a stratified environment, isodiachronic sequential bound estimation was used to spot-check the whale’s position at 15 s intervals starting from the first vocalization. While isodiachronic sequential bound estimation is more accurate than hyperbolic localization, our implementation of this method was computationally intensive and would have taken prohibitively long to analyze every vocalization this way. Performing the sequential bound localization every 15 s served as a quality control check on the hyperbolic localization results.

The same random variables created for the sensitivity analysis were used as inputs into the isodiachronic sequential bound localization algorithm to obtain a cloud of potential whale positions. The shape of this cloud reflects the optimal localization precision and accuracy of the system without requiring the constraint of a homogeneous environment. When localization clouds from the sequential bound estimation are drastically different than those from the hyperbolic localization algorithm, then the assumption of an isovelocity sound speed is likely to be invalid.

E. Trial deployment

On 30 October 2007, the array was deployed around a single male sperm whale diving over the Kaikoura canyon (Fig. 3). The array was deployed from the research vessel Grampus, a 6 m aluminum boat, operating over the Kaikoura canyon with a crew of two. Using the directional hydrophone, three sperm whales were detected vocalizing; however, only one sperm whale was estimated to be within the bounds of the array at the time of deployment. When possible the range and bearing to the diving whale were measured at the surface using a hand bearing compass and laser range finder (Bushnell Yardage Pro Compact 600).

III. RESULTS

Directional hydrophone estimates of the whale’s position indicate that the whale dived near the research vessel and moved toward the southeast. This is consistent with the track generated by the isodiachronic sequential bound estimate (Fig. 4). However, estimates from the directional hydrophone were not precise enough to reveal that the whale descended along the path of a spiral, which can be seen in the 3D tracks computed via sequential bound estimation (Fig. 5). The precision of localization decreased as the whale moved away from the center of the array (Figs. 6 and 7). When localization algorithms could make use of surface reflected multipath the localization precision substantially increased (Fig. 6), which is consistent with results described by Wahlberg et al. The median whale position from the marginal distributions of the hyperbolic sensitivity analysis, $P_x$, $P_y$, and $P_z$, fell within the 95% confidence intervals from the isodiachronic sequential bound analysis (Fig. 7). Maximum depth

FIG. 3. October 30, 2007 array deployment geometry. Bold line shows the track of the boat, while normal lines show the track of the buoys. The thin line with crosses is the whale track, which shown in detail in Fig. 4.

FIG. 4. Joint X-Y whale position probability from isodiachronic sequential Monte Carlo analysis. The whale circles as he dives initially and then heads from northwest to southeast.
was 599 m, while mean depth was 418 m which is comparable to sperm whale diving depths measured in other parts of the world.34,50

**IV. DISCUSSION**

The array has been successfully used to localize diving sperm whales in the Kaikoura canyon. Results and error estimates obtained are consistent with those obtained from other passive sonar systems used to track sperm whales in 3D.17,19,34,45 Tracking whales in three dimensions has the potential to yield information that may not be observable from tagging animals with a depth logger. The spiral at the beginning of the dive in Fig. 5 is a clear illustration of an advantage of 3D tracking. As computer processors, analog-to-digital converters, and digital storage become more powerful and affordable, so should the ability to create inexpensive passive sonar systems with higher fidelity hardware and more sophisticated on-board software.

As computing power increases, it should become feasible to perform the more accurate isodiachronic analysis for every vocalization instead of using it primarily as a quality control check. By using the fast hyperbolic analysis method to localize every click and validating the results with the slower isodiachronic analysis, we attempt to strike a compromise between computation time and accuracy. By using confidence intervals from the sequential bound analysis instead of the hyperbolic analysis, we effectively trade the higher temporal resolution of the hyperbolic analysis for higher overall accuracy of the isodiachronic analysis. This trade-off is only necessary when computing power or analysis time are limited.

Our system does have a few additional limitations. A notable limitation is the limited dynamic range of the recording units, resulting from our choice of inexpensive off-the-shelf field recorders. Because the goal of our system was 3D localization rather than measurement of sonar emission patterns, we view this trade-off as acceptable since it allows detection and localization over greater ranges. Automatic gain control or recording devices with wider dynamic range would address this issue. Another limitation of the system is the amount of time required to process the data from each platform to obtain localization. Presently the largest portion of processing time is spent demodulating the GPS positions, which takes the same amount of time as the duration of the recording. Lastly, a major limitation of the system is that high accuracy localizations only occur when the whale is within the bounds of the array. For whales that can swim several kilometers in a single dive, some luck is required to obtain high accuracy localizations over a full dive cycle. However, even the subset of observations for which the whale remains inside the bounds of the array has the potential to yield important insights into the underwater behavior of these whales.

---

**FIG. 5.** Whale trajectory in 3D from the beginning of the dive (time 00:26:15–00:35:32). Whale spirals as he descends. Each cloud corresponds to one whale vocalization. For clarity successive clouds have different shadings and the solid line connects the median point from each cloud.

**FIG. 6.** The total volume of each localization cloud. Open circles and filled circles show volumes from hyperbolic sensitivity analysis using real hydrophones and using virtual hydrophones from multipath surface reflections, respectively. Squares and diamonds show isodiachron sequential bound analysis again using real and virtual hydrophones. Solid line indicates the precision threshold. Dashed and dot-dashed lines show reference volumes corresponding to spheres with diameters of 45 and 15 m (3 and 1 whale-length), respectively.
FIG. 7. Whale location as a function of time. Dots show the median of the marginal probability distribution from the hyperbolic sensitivity analysis. Squares show the 95% confidence limits from the isodiachronic sequential bound estimation.

The principal advantages of our system are its low cost, portability, and ease of use from a small boat. The instrumentation has no moving parts and can survive bumps and jostles that occur at sea during difficult weather conditions. The system is portable and unlike a fixed hydrophone array, it can be deployed and repositioned around the target animals. Processing occurs on shore and requires a desktop PC with adequate storage space (5–10 Gbytes/recording session). Each one of our buoys can be built using mostly off-the-shelf components with a total cost of the parts under US$1000.

While the main goal of our study is to detect and localize sperm whales in the Kaikoura canyon, the array could potentially be used to localize any loud sound sources in the area including baleen whale vocalizations, shipping traffic, underwater explosions, or construction activity. On one occasion the system was used to localize concurrently not only a nearby diving sperm whale but also a singing humpback whale in Kaikoura at an approximate range of 8 km (B. Miller, unpublished).

Long-term use of the system has the potential to provide insight into whether individual whales have different foraging styles and how diving behavior changes with season. Additionally the system may be used to investigate the effects of anthropogenic noise from sources such as whale watching platforms on sperm whale underwater behavior.

ACKNOWLEDGMENTS

Thanks to Miranda van de Linde, Elanor Hutchison, Liz Slooten, and all of our volunteers for providing help with field work. We thank Hamish Bowman for help with buoy construction and design advice, Ross Vennell for assistance writing localization software, and Paul Denys for many consultations about GPS postprocessing. Bertel Møhl and Niels Kristiansen generously provided their FSK modulator and demodulator designs, as did Aleks Zosuls for bandpass filter designs. Funding for field work was provided by the New Zealand Whale and Dolphin Trust. Funding for instrumentation was provided by Otago University.

21F. Lariviere, “Sperm whale habitat off Kaikoura, New Zealand, a descriptive study of physical and biological processes influencing sperm whale


