Title

A Portable Matched-Field Processing System Using Passive Acoustic Time Synchronization

Journal

IEEE Journal of Oceanic Engineering, 31(3)

ISSN

0364-9059

Authors

Thode, Aaron M
Gerstoft, Peter
Burgess, William C
et al.

Publication Date

2006-07-01

DOI

10.1109/JOE.2006.880431

Peer reviewed

A Portable Matched-Field Processing System Using Passive Acoustic Time Synchronization

Aaron M. Thode, *Member, IEEE*, Peter Gerstoft, William C. Burgess, *Member, IEEE*, Karim G. Sabra, Melania Guerra, M. Dale Stokes, Michael Noad, and Douglas H. Cato

***Abstract*—A portable matched-field processing (MFP) system for tracking marine mammals is presented, constructed by attaching a set of autonomous flash-memory acoustic recorders to a rope to form a four-element vertical array, or "insta-array." The acoustic data are initially time-synchronized by performing a matched-field global inversion using acoustic data from an opportunistic source, and then by exploiting the spatial coherence of the ocean ambient noise background to measure and correct for the relative clock drift between the autonomous recorders. The technique is illustrated by using humpback whale song collected off the eastern Australian coast to synchronize the array, which is then used to track the dive profile of the whale using MFP methods. The ability to deploy autonomous instruments into arbitrary "insta-array" geometries with conventional fishing gear may permit nonintrusive array measurements in regions currently too isolated, expensive, or environmentally hostile for standard acoustic equipment.**

***Index Terms*—Acoustic arrays, acoustic beam steering, marine animals, underwater acoustic arrays, underwater acoustics, underwater technology.**

I. Introduction

TIME synchronization of multichannel underwater acoustic data is a perquisite for any coherent signal-processing technique, ranging in complexity from cross correlation, to plane-wave beamforming, to computationally intensive methods such as matched-field processing (MFP) [1]–[4]. The precision required for time alignment is stringent: Beamforming a 100-Hz acoustic signal requires that the data be synchronized to within a millisecond or less for optimum performance.

The oldest, and still most common, solution to time synchronization is to use connecting wires embedded in a cable to transmit signals from individual hydrophones to a central recording and monitoring site. Some arrays will substitute radio transmissions for physical cable along portions of the signal path. If the data must be monitored in real-time, then there are currently no substitutes for cables, although multiplexing techniques can reduce the number of wires required in the cable.

Array cables, however, add substantial weight to the resulting array system, as well as limit their portability, flexibility, and versatility. The field collection of acoustic array data at sea is thus traditionally a large-scale, large-ship, and large-budget affair, and substantial lead time is generally needed to deploy an array. As a result, many interesting transient ocean phenomena remain difficult to study acoustically. If acoustic data do not need to be analyzed in real-time, as is the case with much ocean acoustic research, then alternative approaches for array time-synchronization become viable.

Here, we present a new modular array concept, whose performance is demonstrated using a recording of a humpback whale song [5]–[12] collected off Queensland, Australia, in October 2003. Instead of a set of hydrophones connected by wires, a set of lightweight, low-power, autonomous acoustic recorders have been attached to a rope to create a 14-m aperture array in 25-m-deep water. The replacement of array cable with rope reduces the anchor weight and subsurface floatation required to maintain the vertical array, reducing the weight of the entire assembly, including anchor, to 12 kg. The array system could thus be deployed and retrieved by hand from a small motorboat. Indeed, without the anchor, the entire system weighs only a couple of kilograms, and is compact enough to transport via carryon luggage on a commercial airliner.

The technology for such a concept has only become viable over the past decade, due to rapid advancements in both the consumer electronic industry and in acoustic monitoring tags for marine mammal and other bioacoustic applications [13]–[15]. The cumulative outcome of these efforts have been palm-size recorders with at least 1 GB of flash memory that weigh less than half a pound, and can be powered with standard AA or AAA batteries.

The fundamental technical hurdle of this concept is how acoustic data individually sampled on each recorder can be time-synchronized to the precision required for coherent processing. The situation is exacerbated by the fact that the internal clocks of extant flash-memory recorders exhibit high drift rates relative to each other—up to 1 s a day. Thus, not only would the signals need to be synchronized, but they would need to be frequently resynchronized.

There are three common approaches for time-synchronizing autonomous recorders: direct, active acoustic, and passive acoustic synchronization. Direct synchronization

Manuscript received October 1, 2004; revised February 3, 2006; accepted March 10, 2006. This work was supported in part by the Office of Naval Research (ONR), Entry Level Acoustic Faculty Award under Grant N00014-03-1-0215 and in part by the Marine Mammals Program under Contract N00014-03-1-0460. Development of the bioacoustic probe was supported by ONR under Contract N00014-99-C-0170. The Humpback Acoustic Research Collaboration (HARC) was supported by ONR Award N00014-02-1-1013. **Associate Editor: J. Preisig**.

A. M. Thode, P. Gerstoft, K. G. Sabra, M. Guerra, and M. D. Stokes are with the Marine Physical Laboratory (MPL), Scripps Institute of Oceanography, San Diego, CA 92106 USA (e-mail: thode@mpl.ucsd.edu).

W. C. Burgess is with Greeneridge Sciences, Inc., Goleta CA 93117 USA.

M. Noad is with the Department of Veterinary Science, University of Queensland, St. Lucia, Qld., Australia.

D. H. Cato is with the Defence Science and Technology Organization, Defence Department of Australia, Sydney, N.S.W., Australia.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JOE.2006.880431

simply means that the recorders are synchronized before and after a deployment, and then a linear clock drift is assumed when synchronizing the data collected during the deployment. This method is standard procedure for synchronizing bottom-mounted autonomous seismic-monitoring packages that are spatially separated by more than a kilometer. However, the assumptions behind direct alignment imply that it cannot synchronize data to the precision required for coherent processing, particularly if the instruments experience substantial temperature changes during the deployment, thus changing the oscillator frequency. In principle, detailed sampling of the clock chip temperature over time might achieve synchronization below 1-ms accuracy if the oscillator's dependency on temperature is sufficiently characterized. However, we are not aware of any success in using direct synchronization to coherently beamform data collected from a set of autonomous instruments in the acoustic-frequency range.

Active acoustic methods are another possible solution [16]. Unfortunately, the high clock drift of commercially available recorders implies that active sources would need to generate signals often, raising concerns about power consumption and weight, particularly in situations where the frequency range of concern is less than 1 kHz. Broadband sources below this frequency tend to posses inconveniently large dimensions and power requirements that obviate many of the advantages to be gained from eliminating interconnecting wires. An even greater disadvantage of active sources is that they are intrusive and subject to regulation, particularly if recordings in the presence of marine life are desired.

Here, two passive acoustic methods are used to time-align the signals on the vertical array, and then track the subsequent clock drift between the instruments. The initial time alignment uses a straightforward extension of global inversion techniques already in use by the acoustics community [3], [17]–[20], while the clock-drift measurements exploit the spatial coherence of the ambient noise field [21]–[26]. Autonomous recorders, geoacoustic inversion techniques, and the coherence characteristics of the ambient ocean noise field are all well-established topics in the literature. However, to our knowledge, these subjects have never been combined into a time-synchronization and tracking system.

Section II of this paper describes the array concept, the recording hardware, and the synchronization procedures. Section III describes the experimental setup that was used to demonstrate the array performance, including the experimental location and deployment scheme. Section IV then demonstrates how a humpback whale song and the ambient noise background were used to synchronize the array data, which were then used to track the whale in range and depth using MFP techniques. Finally, Section V summarizes the issues and implications of this "insta-array" concept, including challenges of extending the technique to other array geometries and larger apertures.

II. Array Description

A. Autonomous Acoustic Recorders

The hardware that makes this array concept possible originated with marine mammal acoustic monitoring tags, originally

Fig. 1. Autonomous instrument package for recording acoustics, pressure, temperature, and inclination data to a 1-GB flash memory chip. The device can be powered by 4-AA batteries (shown here), although 4 AAA were found to be sufficient.

developed to study the effects of anthropogenic noise on individual marine mammal behavior. The acoustic elements used in the present design are slight modifications of a marine mammal tag [13], [14], and illustrated in Fig. 1 with AA batteries inserted. This so-called "bioacoustic probe" was designed to sample acoustic data at sampling rates of 100 Hz to 20 kHz using an HTI-96-MIN/3V hydrophone (typical sensitivity of -172 dB re 1 V/μPa) and storing the data to 1 GB of flash memory with 16-b precision. In addition, auxiliary measurements of pressure, temperature, and acceleration on two axes are sampled once a second and also stored to memory. Four AAA batteries were found to provide sufficient energy to fill the memory. All components except for the hydrophone have been inserted into a transparent acrylic pressure case with a Delrin end-plug, manufactured by Cetacean Research Technology, Seattle, WA. The resulting length and diameter of each recorder is 25 and 5 cm. The hydrophone is connected to the internal electronics via a Subconn underwater connector, and the pressure sensor is embeddeed in an external port of the pressure case. External power can be provided to the electronics via the same connector. The device is programmed via infrared transmissions from either a handheld personal digital assistant (PDA) or a laptop. The accelerometer chip (MXA2500GL, Memsic Inc., North Andover, MA) is mounted perpendicular to the main circuit board to allow the tilt of the recorder from the vertical to be measured along two axes.

A couple of the recorders were also encased in epoxy, instead of using a pressure case, yielding a recorder with even smaller dimensions of 20 cm $\times$ 2.5 cm. This design is powered by a lithium 3.6-V battery.

B. Synchronization Procedure for Vertical Array

The hydrophone synchronization procedure takes place in four stages: 1) coarse correlation of a broadband signal to align the data to within 10 ms; 2) determination of inversion bounds for the remaining time offsets; 3) a fine-scale global inversion, or "focalization," of the acoustic source to time-align the data to within a millisecond or less at a particular "focal" time; and 4) measurements of the ambient noise field to correct for the relative clock drift between each possible hydrophone combination, to permit time-synchronization at times other than the "focal" time. The basic strategy is to initially ignore the fact that the signal has multipath during stage 1 to get a rough time alignment, and then use a detailed acoustic propagation model in stage 3 to explicitly account for multipath and other propagation effects, permitting the extraction of much higher resolution timing information embedded in the signal. To avoid repeating this exercise for every signal of interest, in stage 4 the

ambient noise field can be exploited to correct for subsequent clock drifts.

1) Coarse Correlation: The raw acoustic data downloaded off the phones may be offset in time by up to 2 or 3 s, due to a combination of cumulative clock drift and the fact that the activation time on the recorders is only precise to about 1 s. The first step, therefore, is to roughly align the various time series by cross correlating each pairwise hydrophone combination with each other, during a period of time when a broadband signal is present. The time at which the cross-correlation function between hydrophones i and j attains its maximum value t_{ij} is then selected as the time-shift required to align the data records.

For the particular geometry of a vertical array in shallow-water waveguide, using cross correlation to time-align the data is a crude and coarse technique which, by itself, would be insufficient for achieving coherent processing across the array. Cross correlating two signals implicitly assumes that the signals are propagating in free space and arriving broadside to the sensors, so that in a noiseless environment the received signals at both sensors would be identical.

In reality, a vertical array spanning an acoustic waveguide will receive many multipath arrivals for a given acoustic signal, and thus the signal structure will vary with receiver location. In situations typically explored with MFP methods, this multipath cannot be separated in time, but the various arrivals constructively and destructively interfere such that the amplitude and phase of a given frequency component varies in a complex way with range and depth in the waveguide. At frequencies where the acoustic wavelength is a substantial portion of the waveguide depth, these multipath interference effects can be conveniently modeled as a set of propagating normal modes.

However, by simply cross correlating the two signals, all this signal complexity is blithely ignored and thus potentially useful timing information is neglected. However, the data can be aligned with enough accuracy to subsequently permit a much more detailed propagation model to be used to extract much more precise timing information from the signal. This paper also suggests that the time window used for the cross correlation must be sufficiently long to permit all signal propagation paths to be present in the signal. Thus, the time window should be greater than the time spread expected from geometric dispersion effects (or spread in modal group velocities if propagation is expressed in terms of normal modes).

2) Setting Inversion Boundaries for Clock Offset: Before attempting to invert for more precise timing offsets using a more sophisticated propagation model, some boundaries must be assigned to the inversion search space. These are obtained by measuring the standard deviation of a particular value of t_{ij} from the data, by computing the cross-correlation functions t_{ik} and t_{jk}, where k represents an additional hydrophone. A new estimate of t_{ij} is then derived from the relationship $t_{ij=}t_{ik}-t_{jk}$. Permuting through all possible values of k produces a set of estimates for t_{ij} from which a standard deviation can be computed. Because this method combines cross-correlation measurements between all combinations of phones, and not just adjacent phones, one would expect that the maximum spread in peak correlation times would be on the order of $\pm L/c$, where L is the aperture of the vertical array. That is, in an extreme situation the cross correlation of the signals received by two phones at the array ends could be dominated by vertical multipath, instead of arriving broadside as implicitly assumed in step 1, thus producing a maximum time deviation of L/c from the true clock offset.

For example, in Section III the vertical aperture of the array was 14 m, leading to an expected maximum deviation of the various cross correlations to be $\pm$ 9.3 ms. In reality, the time-offset standard deviation between two phones in the experiment discussed later is about 5 ms. Thus, once the array inclination was checked (as described later) the inversion bounds for the next step were set to $\pm$ 10 ms around t_{ij}.

3) Global Inversion: A variety of global inversion algorithms have been developed to maximize the fit between measured acoustic data and the output of a detailed propagation model [18]–[20], [27]–[29]. Originally developed to determine the best-fit depth and range of an acoustic source in a waveguide, these MFP methods have been extended to search for best-fit solutions for the ocean environment, as well as array shape [16], [30]–[32], a process that has been called "focalization" [17] or "geoacoustic inversion."

Brief Background: There is a large literature on MFP methods and its extensions, so only a brief discussion of the signal processing techniques used in this paper is presented here. Selected references that cover the diverse environments under which these methods have been demonstrated are provided in [2] and [33]–[37]. The performance degradation of matched-field and geoacoustic inversions at low signal-to-noise ratio (SNR) has been a topic of continuing theoretical [38], [39] and experimental [34] research.

Once the data have been roughly time-aligned from step 1, the Fourier transform of a signal sample on each hydrophone is taken. The length of this "snapshot" needs to be long enough to permit a particular signal-frequency component to emerge from the background noise and to include all relevant multipaths, but short enough so that the propagation environment can be considered "frozen" over the time required to collect a sufficient number of snapshots to yield a statistically stable estimate of the cross spectral density matrix (CSDM).

A set of frequency components is chosen from the signal for inversion. For each frequency f_i, the corresponding complex Fourier coefficients from each hydrophone are assembled into a complex column vector $\mathbf{p}(f_i)$. The outer product $\mathbf{p}(f_i)\,\mathbf{p}(f_i)^{\mathbf{H}}$, where the superscript "H" indicates the conjugate transpose of $\mathbf{p}$, is defined as CSDM $\mathbf{K}(f_i)$. The CSDM is normalized by its trace to remove amplitude and phase effects associated with the source waveform, which is typically unknown. What remains is information about the relative phase and amplitude of the signal component between each hydrophone pair, thus preserving any signal multipath effects. Another advantage of converting the data to a CSDM is that subsequent snapshots can be averaged together to improve the data SNR, provided that the source location or propagation environment do not change appreciably during the averaging time. The ratio of the largest eigenvalue of the CSDM to the sum of the remaining eigenvalues provides an indication as to whether the propagation environment and source location are sufficiently stable to permit an inversion

[37]. A large value of this ratio indicates that the signal structure across the array aperture, and thus the propagation environment, is consistent throughout the averaging process.

After the CSDM is obtained, an acoustic propagation model with the ability to explicitly model multipath and other waveguide propagation effects is then used to model the complex pressure field measured by the array from a source located at a candidate location r and z. This "replica" vector $\mathbf{r}(r, z, f_i)$ can be combined with the data CSDM in a variety of ways, but the simplest and most frequently used is the Bartlett beamformer

$$B(r, z, f_i) = \frac{\mathbf{r}(r, z, f_i)^{\mathbf{H}} \mathbf{K}(f_i) \mathbf{r}(r, z, f_i)}{|\mathbf{r}(r, z, f_i)|^2}. \tag{1}$$

When computed as a function of modeled source range and depth, the output of (1) generates an "ambiguity surface" with values that vary between 0 and 1, with 1 indicating a perfect match between model and data. The entire procedure can be repeated at different frequencies and the resulting ambiguity surfaces can be incoherently averaged together to suppress locations that generate local maxima or "sidelobes." In controlled field experiments in the open ocean, frequency-averaged Bartlett values greater than 0.8 are generally considered a good result (e.g., [2]), but much lower values can also give stable results (e.g., [34]).

The concept behind "focalization" is simply to expand the parameter space beyond source range and depth to environmental parameters and receiving array shape [17]. As the modeled environment and modeled array geometry approach to the actual propagation environment and array geometry, the output of (1) should approach 1. Thus, with a good acoustic propagation model, one can exploit the multipath present in the signal to achieve much better time-alignment than was possible with simple cross-correlation techniques.

Modification for High-Resolution Synchronization: The ability to invert acoustic data for array shape is relevant to the effort discussed here, because small clock offsets between the data can be interpreted as small horizontal displacements of the hydrophones relative to each other, for the case of an acoustic signal propagating in a waveguide. Therefore, if enough multipath information exists in the propagating acoustic signal to invert for array shape, intuitively one might expect that clock offsets could be inverted as well.

In the work presented here, the normal mode propagation model of the genetic algorithm inversion code seismoacoustic inversion using genetic algorithms (SAGA) [40] was modified to incorporate relative clock offsets into a geoacoustic inversion. The SNAP normal mode model was used because the frequency bandwidth of the data of interest indicated that the corresponding acoustic wavelengths would be on the order of the waveguide depth, thus making the normal mode formulation a more accurate representation than ray-tracing procedures for this environment. SAGA was modified to incorporate additional parameters δ_j, or the clock offset of the jth hydrophone relative to the shallowest hydrophone. For a given frequency f, the modeled pressure field at the jth hydrophone was then multiplied by the phase factor $e^{i2\pi f \delta_j}$ to permit comparison to the data via (1). The limits of the inversion span for the jth hydrophone's clock offset were derived from the standard deviation of t_{1j} computed from step 2. In the acoustic inversions presented here, each individual clock offset was allowed to vary between ± 10 ms from the reference hydrophone, based on the results of stage 2 and the autonomous recorder inclinometer measurements.

A natural concern is whether more than one set of relative clock offsets could produce the same fit to the acoustic data—i.e., does solving for clock drift make the problem underdetermined, particularly if an array with few elements is used? No single analytic argument is persuasive, but in Section IV inversions of both modeled and real data indicate that a global inversion can be achieved even on a sparse array, provided that a sufficient number of frequencies are incorporated into the inversion, and that the inversion boundaries are not too large. Inversions of modeled scenarios, in particular, provide a useful insight into the number of acoustic-frequency components required to obtain a unique solution. Care must also be taken to avoid selecting acoustic frequencies that have a harmonic relationship, which permits multiple candidate clock offsets to yield the same relative phases for each modeled frequency component, creating a nonunique solution. These offsets lie at integer multiples of the period of the lowest frequency component. Therefore, the frequencies selected for inversion should not have ratios that can be expressed as ratios of small integers. A good strategy is to select a few frequency components over a narrow frequency range, and then to select the rest over a broad frequency range, a strategy similar to that used by multiple frequency continuous wave (CW) radar to eliminate range ambiguities [41].

A final concern about this synchronization approach is how to treat substantial physical array tilt, because as mentioned previously, small physical lateral offsets generated by a tilted vertical array are analogous to clock offsets arising from time misalignment, particularly whenever a clock offset is a small fraction of the period of the frequency in question. This ambiguity has no effect on the inverted source location. For a slightly tilted array, auxiliary inclinometer information can be used to separate relative phase differences arising from lateral element offsets from phase differences arising from clock offsets. What comprises a "nearly vertical" array depends on the total array aperture and maximum acoustic frequency of interest. A 14-m aperture array receiving a 1-kHz signal would need to have a vertical inclination of less than 3° to keep the lateral hydrophone offsets to less than 10% of an acoustic wavelength. We will find in the following data that a vertical array tilt of up to 5° could be neglected, while still achieving acceptable MFP resolution, provided that data from a broadband of nonharmonic frequencies is incorporated. In Conclusion, possibilities for handling a substantially titled array are discussed.

4) Clock-Drift Estimate Using Background Noise: The previous inversion procedure aligns the autonomous recorder data at a single instant, but given the limitations of current recorder crystal oscillators, the data will desynchronize by 1 ms within a minute of original "focal" time. In principle, one can simply perform a series of inversions to estimate how the clock offsets evolve over time, but this is computationally expensive. A more

convenient approach is to exploit the spatial coherence of the ambient noise background [21], [22], [25], [26], [42]–[46] to estimate the relative clock drifts between instruments.

The "background" ambient noise recorded on two hydrophones is spatially correlated, with the strength of the correlation decreasing with increasing instrument separation. This correlation arises because the background noise signal is the cumulative sum of a large number of individual source events, each of which propagates through the same waveguide environment. Sometimes the propagation path from a particular acoustic event passes through both hydrophones, and if enough of these events occur over a given time interval, the ambient noise between both phones will become correlated. Several analytical expressions for the ocean ambient noise correlation have been derived, including one for an infinitely deep ocean [21] and for a range-independent waveguide [22]. Recent work on these expressions indicates that an estimate of the Green's function connecting the hydrophones can be extracted from samples of ambient noise [43]–[48].

Here, three properties about the ambient noise field are assumed. First, the noise between hydrophones is assumed to be spatially correlated, so that a nonzero normalized mutual coherence function can be measured in the frequency domain via

$$\gamma_{ij}^{T_0}(f) = \left\langle \frac{X_i^{*T_0}(f)\, X_j^{T_0}(f)}{\left|X_i^{T_0}(f)\right| \left|X_j^{T_0}(f)\right|} \right\rangle \tag{2}$$

where $X_j^{T_0}(f)$ is the discrete Fourier transform of a segment of data recorded on hydrophone j at a given time T_0, evaluated at frequency f. The asterisk designates complex conjugation, and the brackets $\langle\rangle$ represent the expectation operator, or ensemble average. The inverse Fourier transform of (2) will be defined as the normalized coherence time function, or "coherence function" for brevity [49]. The second assumption is that the ambient noise statistics are ergodic, so that the expectation operation in (2) can be interpreted as a time average over a short-time interval. Finally, in the absence of clock drift, strict-sense stationarity is assumed in that a measurement of the normalized coherence function (2) at time T_1 would be equal to the value measured during T_0. That is, both the statistical properties of the ambient noise sources and the characteristics of the propagation environment are assumed to remain relatively constant between T_0 and T_1.

Suppose that (2) is computed at a time T_0 where the data have been synchronized, and then recomputed at a later time T_1 when the data have become misaligned in time by an amount τ. We assume that the clock-drift rate is small enough that a short data sample subjected to a Fourier transform experiences no effective drift within that sample. From the Fourier shift theorem, one obtains

$$\begin{aligned}\gamma_{ij}^{T_1}(f) &= \left\langle \frac{X_i^{*T_1}(f)\, X_j^{T_1}(f)}{\left|X_i^{T_1}(f)\right| \left|X_j^{T_1}(f)\right|} \right\rangle \\ &\approx e^{i2\pi f\tau} \left\langle \frac{X_i^{*T_0}(f)\, X_j^{T_0}(f)}{\left|X_i^{T_0}(f)\right| \left|X_j^{T_0}(f)\right|} \right\rangle = e^{i2\pi f\tau} \gamma_{ij}^{T_0}(f)\end{aligned} \tag{3}$$

where the third assumption has been employed in the third expression. The resulting normalized coherence function $\gamma_{ij}^{T_1}(t)$ simply becomes a time-shifted version of the original coherence function measured at T_0

$$\gamma_{ij}^{T_1}(t) = \gamma_{ij}^{T_0}(t-\tau)\,. \tag{4}$$

The implication of (3) is that the clock drift τ can be measured simply by tracking how a stable peak of the coherence function shifts with time, even if the drift is nonlinear in time. Equations (2) and (3) also assume that the time segments collected from the two hydrophones have at least some time overlap, and that the relative separation between the hydrophone pair remains constant between T_0 and T_1. The auxiliary pressure, inclination, and temperature measurements on the autonomous recorders can be used to confirm or adjust this last assumption.

After extracting a clock-drift estimate from the normalized coherence time function, the acoustic data can be time-aligned without performing a computationally intensive inversion for every event of interest. While the clock-drift estimates used here were manually extracted from a plot of the coherence function verus time, automating this measurement should be possible.

III. Experimental Setup

A. Location

Between October 17 and 25, 2003, acoustic data to test the concept were collected off Peregian Beach in Queensland, Australia, about 150 km north of Brisbane. For several years this area has been used to acoustically monitor the southerly spring migration of humpback whales [5], [7]–[9]. During this season, anywhere from 10 to 40 groups of whales, or "pods," swam by the observation site per day, of which around 12% would have singing animals [7]. By 2003, an international collaborative effort, the Humpback Acoustic Research Collaboration (HARC), had been formed to perform an integrated visual and acoustic study of the animals' migration behavior over a two-year period. Between September 7 and October 31, 2003, acoustic data were collected and monitored in real time using a distributed horizontal array of moored hydrophones, each linked by cable to a surface buoy that radioed the data to a shore station. The positions of the hydrophones were fixed by the moorings and accurately determined by theodolite sightings from shore. Localization of singing animals in range and azimuth were made using standard hyperbolic techniques [50], [51]. The moored hydrophones were arranged in a "T" shape with a linear aperture of roughly 2 km a side (Fig. 2), resulting in a system with less than 5% tracking error for sources closer than 2-km range [51]. A nearby hill, Mount Emu, provided a convenient location for concurrent visual observations and theodolite fixes out to 10-km range.

B. Experimental Deployment

During the 2003 HARC field season, the mobile vertical array was deployed in approximately 25-m-deep water, near the tip of the "T" formed by the moored hydrophones (Fig. 2), at roughly 26.5° S, 153.2° E.

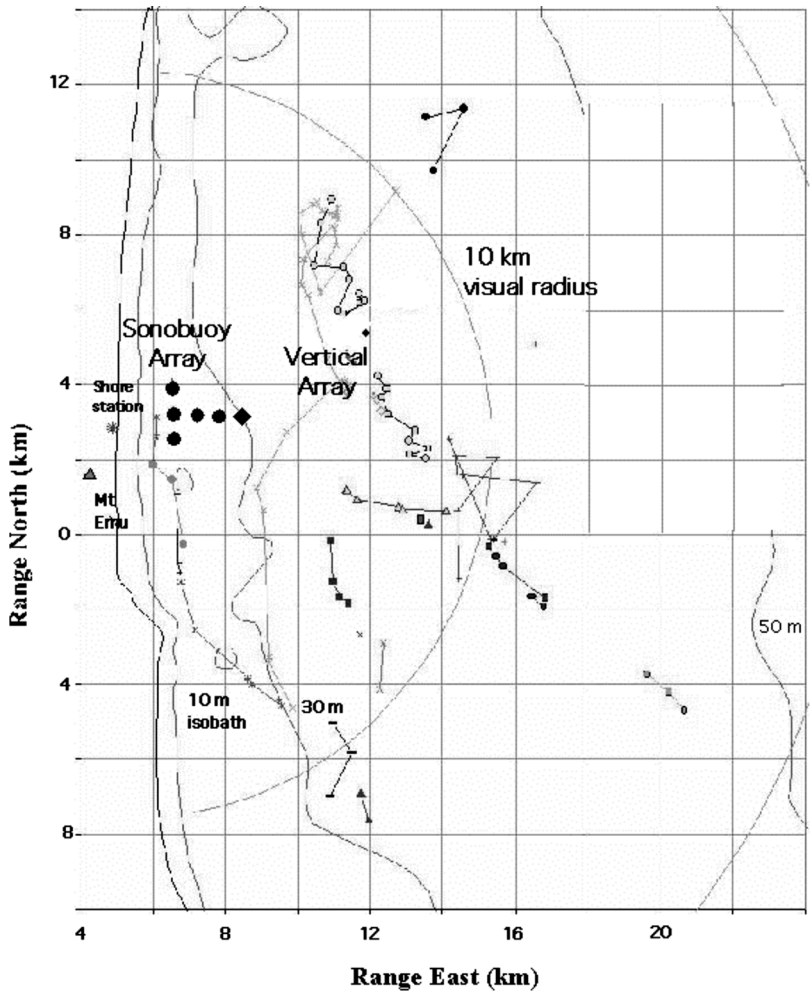


Fig. 2. Relative position of vertical array (diamond) with respect to anchored moored hydrophone array (circles) and coast. Mount Emu is the visual observation site for HARC. Examples of acoustic tracking from the moored hydrophone array are also visible.

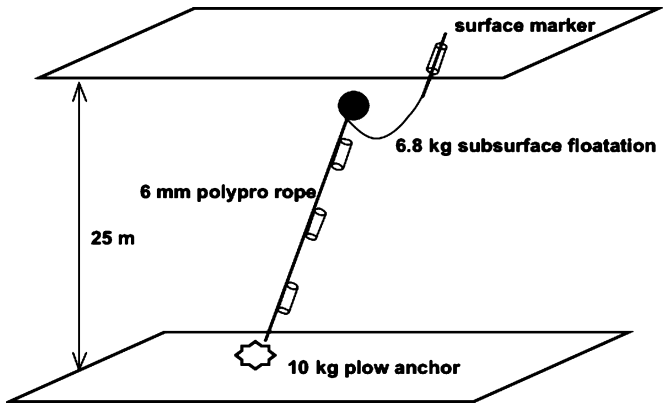


Fig. 3. Schematic of array deployment geometry.

The array consisted of a standard 10-kg plow anchor attached to about 1 m of 0.25-in anchor chain, tied into a 0.5-in braided polypropylene rope (Fig. 3). The other end of the rope was attached to about 5 kg of subsurface floatation in the form a standard plastic fishing float. A short length of bungee connected the float to a surface marker. The currents in the area were found to be mild, so the resulting vertical array tilt was subsequently determined to be less than 5°, once biases in the inclinometers were accounted for. The total array weight, including anchor, was about 16 kg, and thus easily fit into a small motorboat used for biopsy and tagging work. The gear could be deployed and retrieved by hand without much difficulty.

The autonomous recorder pressure cases were taped directly to the rope, but the hydrophones were not taped to the rope, an arrangement that fortuitously minimized pickup of cable strum. Three deployments were conducted, each with a different element spacing; on October 23, 2003, the element spacing was 2.75 m, with the shallowest element at 7.32-m depth. The pressure cases were attached to the rope so that the accelerometer measured in-plane tilt and out-plane tilt. The acoustic data were sampled at 2.048 kHz, which allowed the instruments to record for over 71 h before filling the 1-GB memory.

Acoustic observers constantly monitored the moored hydrophone signals. Whenever humpback song was detected within 5 km of the vertical array position, the acoustic observers made several hyperbolic fixes to provide an independent measure of the range of the animal from the vertical array. Theodolite measurements of surface blows measured by visual observers from Mount Emu were wirelessly transmitted to a base station that logged visual sightings and acoustic tracks into one database.

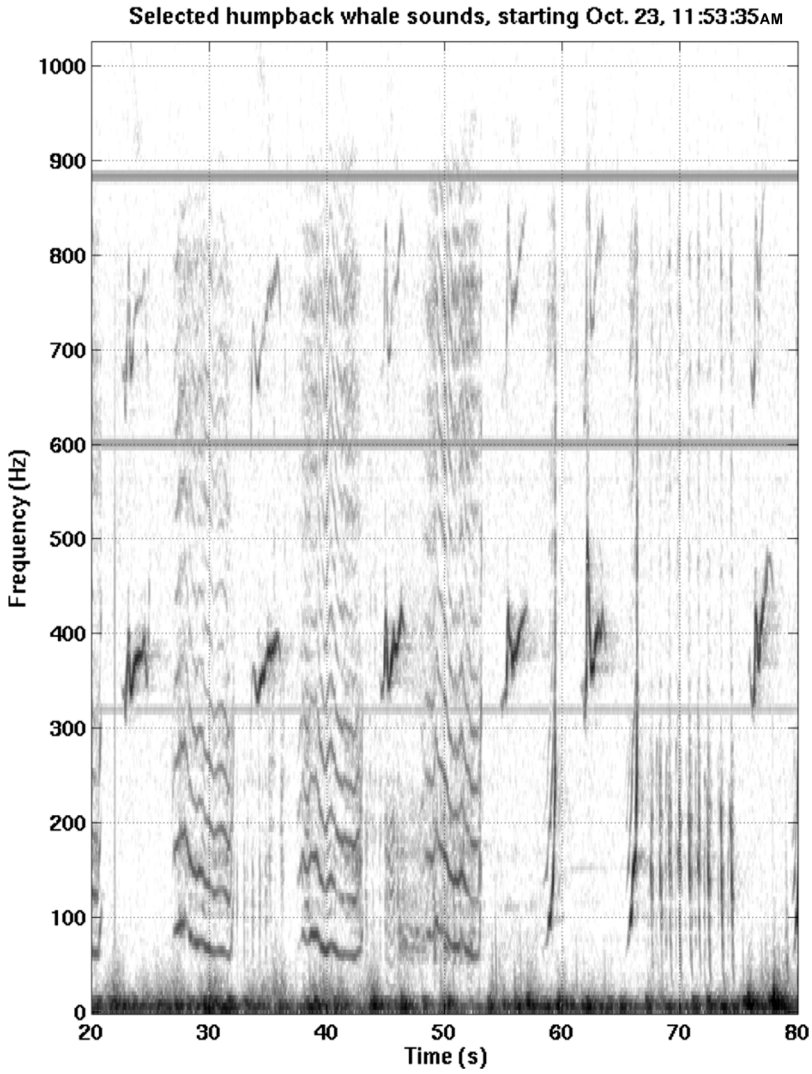


Fig. 4. Spectrogram of humpback whale song used in matched-field synchronization and tracking analysis. The highly modulated signals display significant energy content between 50–900 Hz. Horizontal lines are artifacts of the recording electronics. Calls with many harmonics (e.g., 30 s) and rapid upsweeps (e.g., 60 s) were found to be particularly effective for inversion and tracking.

IV. Experimental Results

A. Signal Characteristics

Between October 17 and 22, 2003, several singing whales passed relatively close to the vertical array site. The closest approach by an individual took place between 11:45:00 A.M. and 12:20:00 P.M. on October 23, 2003.

A humpback song has a hierarchical temporal structure wherein 2–3 distinctive sound "units" are associated together as a "phrase." These phrases are repeated in certain patterns that are called "themes," and these themes are cycled through in an consistent pattern to form a "song" [6], [9]–[11], [52]–[54]. A representative selection of sound units shown in a transition between theme "A" and "C" of an eastern Australian song is illustrated in Fig. 4. Several hyperbolic fixes throughout the animal song indicated that the animal range at one point was less than 300 m from the array. This whale, labeled "D" in the HARC records (fourth active singer of the day), seemed to provide a good candidate for testing of the passive acoustic synchronization method, although only four hydrophones were still recording during this time.

Humpback whales produce a large variety of sound units, but Fig. 4 shows that a majority are highly frequency-modulated (FM) sweeps, which last between two to six seconds per unit. The signal modulation distributed the acoustic energy over many frequency bins, and it was usually possible to find nonharmonic frequency bins with significant energy. For any given song unit, a CSDM was formed for each frequency bin by averaging the Hanning-windowed data throughout the call duration using a 75%-data overlap and 1024-pt fast Fourier transform (FFT) size. This FFT length corresponds to 0.5 s, which is of long enough duration to encompass all expected multipath arrivals from signals produced from sources up to a few kilometers range.

Measurements of the phase coherence of acoustic signals propagating through highly variable shallow-water environments have found that up to 1 kHz signals can be expected to remain self-coherent over a 10–60-s interval, and often longer

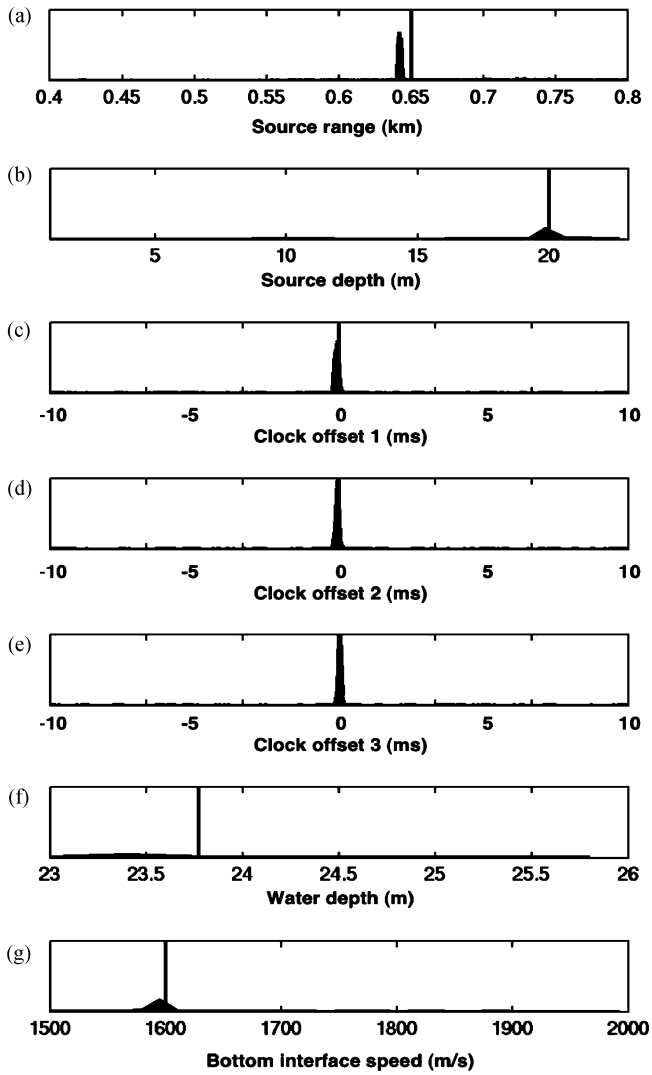


Fig. 5. Histogram of inversion results of 20 SAGA runs on simulated signal at 650-m range and 20-m depth, detected over a four-element vertical array. Key inversion parameters are as follows: (a) inverted range, (b) depth, (c–e) clock offsets of the three bottom hydrophones with respect to the shallowest hydrophone, (f) ocean depth, and (g) bottom interface sound speed.

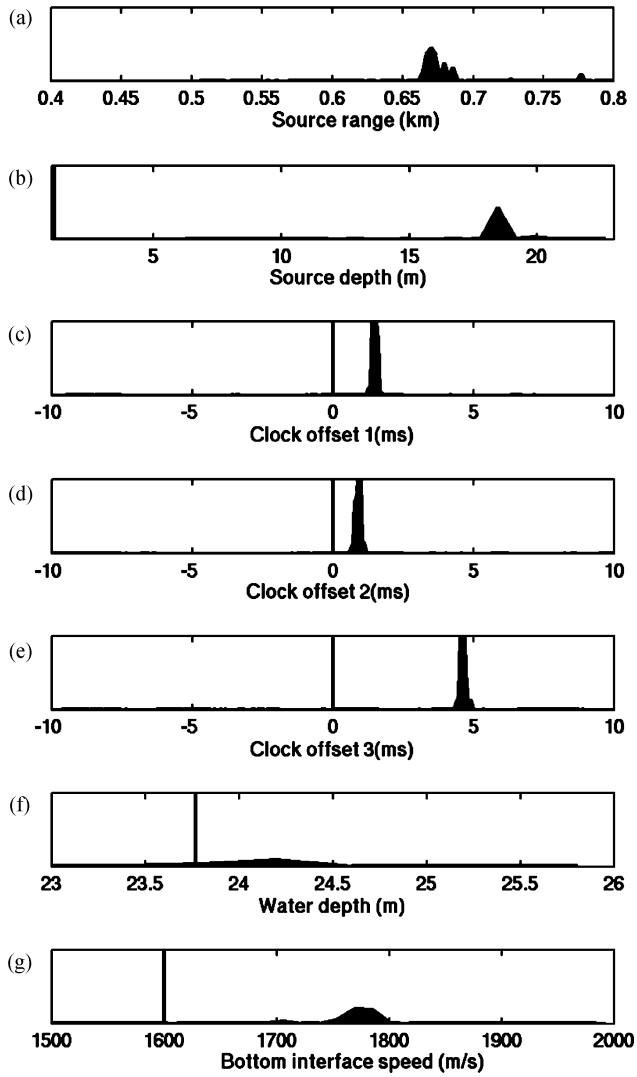


Fig. 6. Histogram of inversion results of 20 SAGA runs for humpback song unit recorded at 11:52:51 A.M., on October 23, 2003. Key inversion parameters are as follows: (a) inverted whale range, (b) depth, (c–e) clock offsets of the three bottom hydrophones with respect to the shallowest hydrophone, (f) ocean depth, and (g) bottom interface sound speed.

[55]. As humpback calls typically last on the order of a few seconds, the animal position and propagation environment were expected to be static throughout the call.

This assumption of a "frozen" environment was verified by computing the ratio of the first singular value of the averaged CSDM to the sum of the remaining eigenvalues. If a CSDM were formed by averaging data from a rapidly fluctuating environment, then the eigenvalues of the CSDM would tend to equalize and the ratio would approach unity. However, primary eigenvalues from a CSDM containing the song units typically produced ratios between 10–25 dB, indicating that the received signal structure was well represented by a single source propagating through a static environment. Typical signal segments of 1–2 s were processed, which provided between 13–29 CSDM snapshots using 512-pt FFT samples with 75% overlap.

B. Initial Inversion and Inversion Simulations

A modulated FM sweep at 11:52:51A.M., produced just a few minutes after whale D began singing, was chosen as the source for the initial time-offset inversion. The coarse cross-correlation procedure found that the acoustic records were initially misaligned by as much as 2.2 s. The standard deviation in the correlation peak times was found to be about 4 ms. At this time, the inclinometers on each element measured vertical array tilt magnitudes averaging less than 5°, a value that remained steady throughout the encounter. Over a vertical aperture of 14 m this tilt could conceivably produce a physical lateral offset of 5 m between the shallow and deepest hydrophone. Thus, the combined effects of array tilt and clock offset could conceivably yield effective time offsets between ± 10 ms.

Ten frequencies spaced between 100–800 Hz were selected from the song unit. Each frequency had at least a 15-dB estimated SNR, and the ratios of each frequency combination could

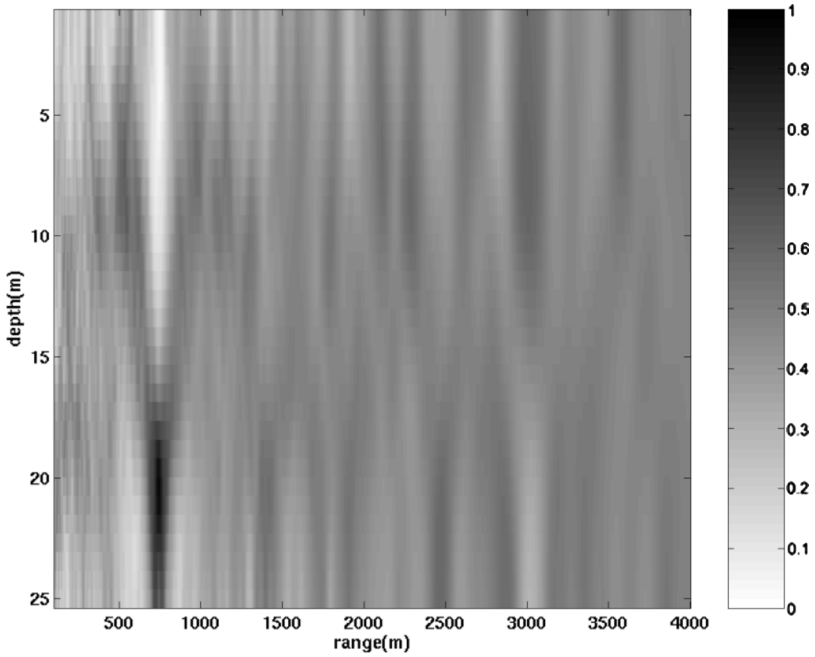


Fig. 7. Frequency-averaged Bartlett MFP ambiguity surface produced using best-fit environment and clock offsets obtained from the SAGA inversion. The maximum Bartlett value is 0.92 for a source range of 661 m and depth of 19.8 m. Ten frequencies, roughly evenly spaced between 100–800 Hz, were incoherently averaged to produce this plot.

not be expressed as a simple rational fraction, thus minimizing the odds of inadvertently obtaining an incorrect clock offset.

The acoustic genetic algorithm inversion code SAGA was used to compute a Bartlett ambiguity surface [24] for each frequency, after which all surfaces were incoherently averaged together, since the absolute amplitude and phase of the song characteristics were unknown. The global inversion was run 20 times, each time performing 4000 iterations on a population size of 64 individuals, and then ending with a local optimization on the best result using Powell's method [40]. The inversion parameters included the three unknown clock offsets, source range and depth, water depth, bottom composition, and sound-speed profile. The last was inverted by generating five empirical orthogonal functions [56] (EOFs) from a set of conductivity–temperature–depth (CTD) measurements made throughout the HARC observational period.

As the array contained only four phones, there were concerns that nonunique solutions to the inversions would exist, and that the global inversion would converge to incorrect source locations, environments, and/or clock offsets. To address this potential objection, inversions of simulated scenarios were performed using the same ten frequency components present in the FM sweep, for representative acoustic environments. The acoustic field produced by a source placed at the data-inverted range and depth was modeled using a normal mode code, and simulated CSDMs were produced by combining the modeled received field with spatially uncorrelated white noise with an SNR of 15 dB. The inversion procedure was then run on the simulated CSDMs to determine whether the inverted parameters were the same as the original modeled parameters. If there had been insufficient information in the received acoustic field to simultaneously resolve source position and array clock alignment, the global inversion would have identified numerous parameter combinations that yield large values of the ambiguity function.

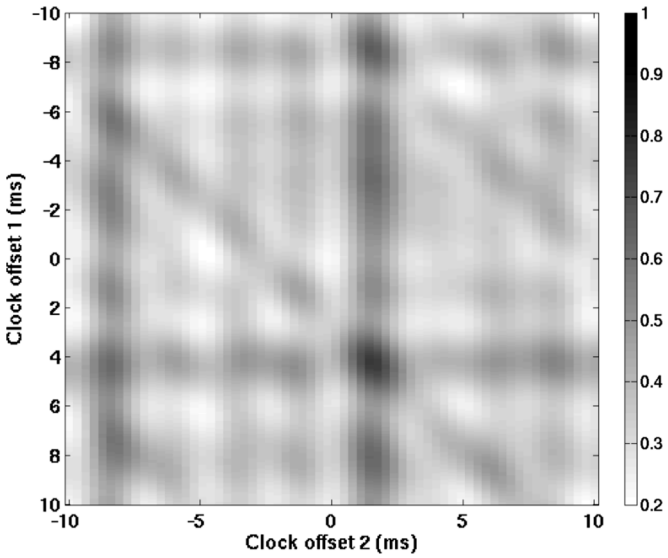


Fig. 8. Bartlett ambiguity surface of 11:52:51 A.M. whale call computed along individual clock offset values for hydrophones 1 and 2, with all other optimized parameters held fixed. The horizontal separation between the mainlobe and the horizontal sidelobe corresponds to the period of the lowest frequency component in the ambiguity surface.

Fig. 5 shows one result of 20 simulated inversions using a source modeled at 650-m range and 20-m depth. For each parameter, the 20 inverted parameter values are plotted as a histogram, with the x-axis limits set to the inversion range of each parameter. A narrow vertical line indicates the true value of the modeled parameter. It is clear that all 20 inversions converge to the correct initial clock offsets of 0, the modeled source range

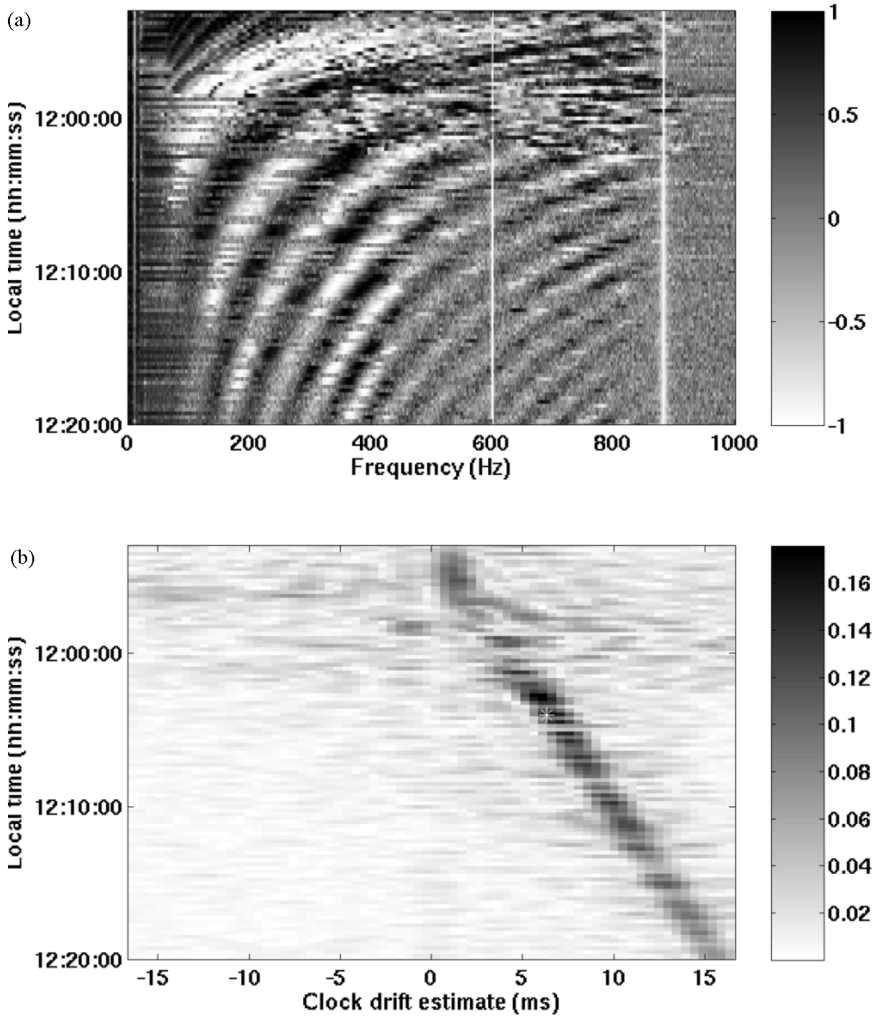


Fig. 9. (a) Normalized spatial coherence as a function of frequency and local time. (b) Inverse Fourier transform of (a) along horizontal axis. The horizontal axis is in units of milliseconds of drift. Plots were generated by averaging 10 s worth of data every 15 s over a 50-min time interval. An FFT size of 1024 pt with 75% overlap was used. The shift of the prominent streak in (b) provides a measurement of the relative clock drift for this hydrophone pair. The white star shows the clock offset derived from a second geoacoustic inversion at 12:04:00 P.M.

and depth, and the correct bottom interface speed. To test the robustness of the timing inversion, the inversion bounds for source range were expanded to 100 m and 10 km, and ± 17 ms for the clock offsets. Exploring this large parameter space proved challenging to the genetic algorithm, but 1 of the 20 inversion runs was still able to converge to the correct model parameters, when the genetic algorithm population size was set to 128. The other inversion runs converged to local minima that were suboptimal, with Bartlett powers of only 0.6 compared with the value of 0.92 obtained at the true solution. Local minima were identified at 3-km source range and clock offsets of ± 10 ms, which correspond to the acoustic period of the lowest frequency component of the inversion.

These simulation results demonstrate that with a careful selection of genetic algorithm parameters, and with enough inversion runs, standard geoacoustic inversion methods can locate the global maximum of the ambiguity function, and thus obtain the correct timing offsets, even when confronted with a large parameter search space. Furthermore, these simulations demonstrate that the introduction of individual clock offsets as inversion parameters did not create an underdetermined inversion problem—only the correct solution produced a significant peak in the ambiguity function, and all local optima identified by the inversion resulted in much poorer fits to the data.

Having demonstrated that this propagation environment permits a robust global inversion for clock offsets, the same SAGA procedure used in the simulations was then applied to the data from 11:52:51 A.M. The bounds on the acoustic source range were initially set between 50 m and 3 km, which was the maximum expected tracking range of the array. A second inversion over a narrower source range between 400–800 m then followed, which permitted fine-scale sampling of the range grid (about 3-m horizontal grid spacing). The water depth was assumed to be independent of range, which is generally a good assumption for this area, although the water depth does vary by a meter or two in certain regions, a fact that becomes relevant in later discussion.

Fig. 6 plots a portion of the subsequent narrow-range inversion results from the humpback song data, in a manner identical to the simulation results in Fig. 5. The histograms derived from

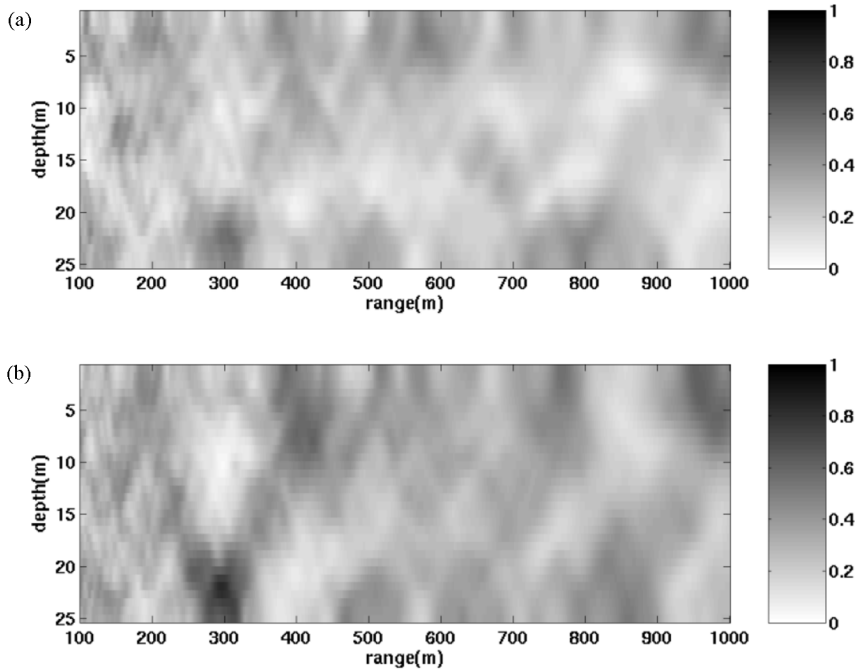


Fig. 10. Effect of clock drift on MFP tracking. (a) MFP ambiguity surface for a sound unit recorded at 11:56:55 A.M., generated using the inverted propagation environment from the 11:52:51 A.M. sound unit. The inverted clock offsets have been held fixed. Maximum Bartlett correlation, averaged over six frequencies, is 0.57. (b) New ambiguity surface generated by updating the inverted clock offsets with the clock-drift information estimated from the ambient noise data. Maximum Bartlett correlation is now 0.84.

data indicate that the crucial model parameters are source range and depth, as well as the relative clock offsets between the hydrophones, just as predicted by the simulations. There is no evidence that an alternative arrangement of clock offsets produce a modeled solution that resembles the measured data, as predicted by the simulations. The inversion also produces a fairly consistent measurement of the bottom interface sound speed of around 1770 m/s. The other inversion parameters, such as subsurface bottom speed and water-column sound speed, tend not to converge to a consistent value, which indicates that the received acoustic field was relatively insensitive to these environmental parameters.

The Bartlett ambiguity surface produced by the best-fit environment and clock offsets is shown in Fig. 7. At a range of 661 m and depth of 19.8 m, the mean normalized Bartlett power is 0.92, averaged over ten frequencies. The width of the mainlobe, defined as the point where the Bartlett power in (1) drops below 0.75 is ± 20 m in range and ± 2 m in depth. This provides a rough experimental measure of the precision of the tracking method. However, the accuracy of the MFP method may be less if the depth of the ocean waveguide is modeled inaccurately due to "mirage" effects [57].

In Fig. 8, the Bartlett surface has been computed as a function of the individual clock offsets to demonstrate that only one set of clock offsets (1.5 and 0.93 ms) produces a global solution. However, it can be seen that when one of the clock offsets reaches -10 ms, or the acoustic wavelength of the lowest frequency component of the inverted signal (104 Hz), a strong sidelobe appears. These results emphasize the importance of selecting frequency components whose ratios are not rational numbers. This plot also demonstrates that the offset ranges are not strongly coupled, explaining in apart why the inversion algorithm consistently converges to the tight distributions visible in the individual offset histograms in Figs. 5 and 6.

C. Clock-Drift Recovery Using Ambient Noise

Having determined the initial clock offsets required to synchronize the data at 11:52:51 A.M., the next goal was to determine the relative clock drift between the autonomous recorders using the spatial coherence of the ambient noise. Fig. 9 displays the measured coherence function defined in (2) in both the frequency and time domain for the two hydrophones in the middle of the water column, which were separated by 2.4 m. By tracing the shift in the prominent peak in Fig. 9(b) over time, one finds that the relative clock drift for this pair of phones is 0.55 ms/min or 33.2 ms/h. The clock drift derived by obtaining updated clock offsets from a second geoacoustic inversion at 12:04:00 A.M. is plotted as a circle in Fig. 9(b), and shows excellent agreement with the ambient noise prediction. Similar agreement was obtained for other hydrophone pairs.

These measured clock drifts, beginning at 11:52:00 A.M., allow the effective clock offsets of the hydrophones to be computed at other times without the need to resort to another full-field inversion. The necessity of correcting clock drift for coherent processing is illustrated in Fig. 10, which shows the Bartlett MFP ambiguity surface for a humpback whale sound unit recorded nearly 4 min after the sound unit used for the initial inversion in Section IV-B. Without correcting for clock drift, the array swiftly "defocuses" and the animal position is lost.

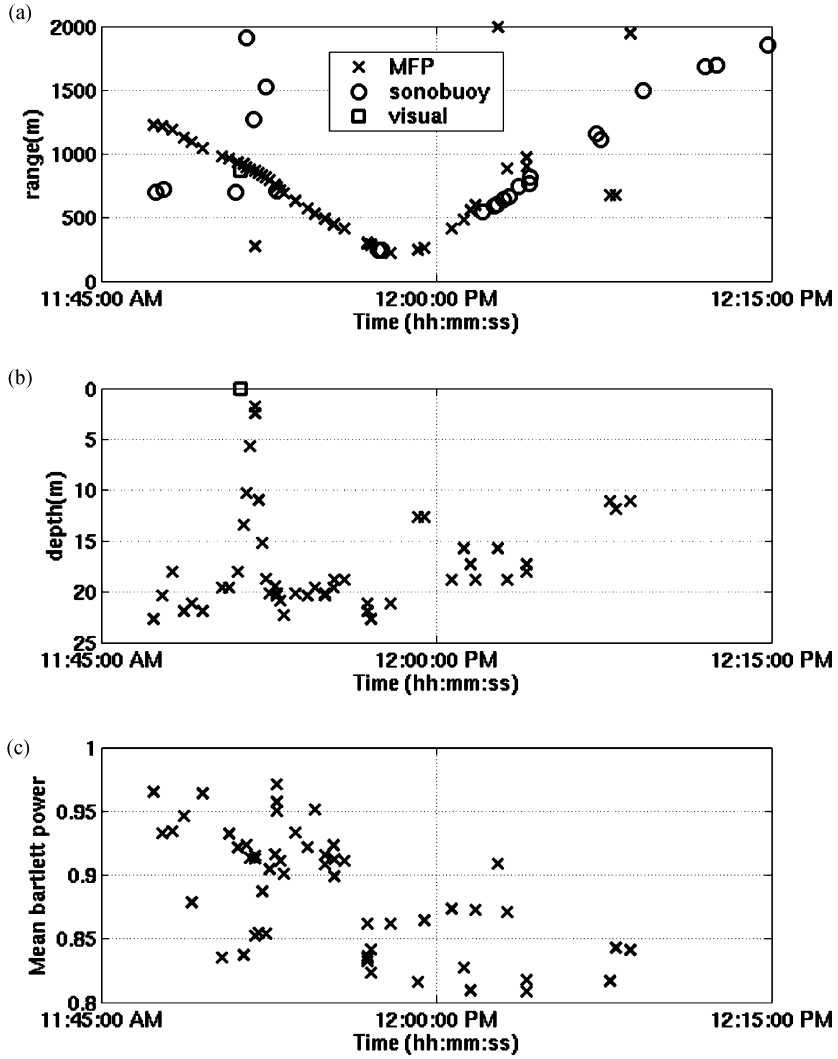


Fig. 11. MFP track of humpback whale using best-fit propagation environment derived using data at 11:52:51 A.M., and using clock-drift measurements from ambient noise to resynchronize data. (a) Range of whale in meters from array site. The "x" represents an MFP location and "o" is an independent hyperbolic fix using the moored hydrophone array, and "□" represents a visual sighting of a surfacing animal. (b) Whale depth versus time. A surfacing of the animal at roughly 11:52:00 A.M. is visible, associated with a visual sighting. (c) Mean Bartlett power averaged over frequency. Note the drop in correlation after the animal swims by the array.

D. MFP Range-Depth Track of Humpback Whale

Having time-aligned the acoustic data to within a millisecond, and having estimated the subsequent clock drift, a number of sound units from animal D were then selected and subjected to MFP processing without conducting further inversions [3]. A minimum of four frequencies was selected for each call, and the eigenvector associated with the largest eigenvalue of each CSDM was used as the data vector for the MFP algorithm. The mainlobe position from each frequency-averaged ambiguity surface was used to produce the tracking plot in Fig. 11. During the time in question the animal was moving south.

Several interesting features can be noted on this plot. First, the MFP ranges derived from the vertical array closely match the ranges obtained from hyperbolic fixes using the distributed moored hydrophone array, as well as a linked visual sighting of a surfacing animal measured by theodolite. The divergence between the two acoustic ranging methods after 12:02:00 P.M. is not due to timing uncertainties between recorders, but is a well-documented feature of the MFP tracking procedure for situations where the bottom bathymetry is not modeled accurately [57], [58]. For example if the bottom bathymetry in the MFP propagation model is 1 m shallower than the average depth of the true bathymetry profile between the whale and array, then the range estimate will be 8% greater than the range obtained from the large-aperture horizontal array. A high-resolution bathymetry map shows that although the water depth at the array deployment site was 25 m, the area 1 km south/south-east of the site deepens to 27 m, and thus the discrepancy between tracking methods after 12:02:00 P.M. is explained.

Second, the humpback whale seems to swim only a couple of meters above the ocean floor, except for a brief period when the animals surfaces at 11:52:00 A.M. Third, the animal seems to continue calling during ascent and descent, hardly pausing to take a presumed breath at the surface before resuming its song. Finally, there is a marked drop in correlation once the animal swims past the array at 11:59:00 A.M. Much of this drop seems due to a decrease in both the sound units' power spectral

density and SNR after 12:00:00 P.M., although potential directivity effects in the humpback signal above 500 Hz cannot be discounted.

V. Conclusion

A modular vertical array has been assembled, by attaching a set of autonomous recorders to a rope. The elimination of interconnecting wires results in a lightweight, robust system that can be deployed in multiple configurations from very small vessels while consuming little power. The system tradeoff is that the data cannot be monitored in real time, and that more computational effort must be expended to synchronize the recorded acoustic data. The results of this paper show that for the particular geometry of a vertical array in a shallow ocean waveguide, sufficient information exists in both a marine mammal acoustic signal and the ambient noise background to synchronize the data, thus permitting MFP, a demanding coherent array processing algorithm, to be employed. This result is only the second peer-reviewed result of applying coherent MFP methods to a marine mammal [3], and would not have been logistically possible without using this particular system.

Many questions remain about the practical performance of this array concept. One immediate concern is how to extend the method to situations where the vertical array tilt is large enough that the physical lateral offsets of the hydrophones must be explicitly modeled. Under this circumstance, the procedure presented in Section II would still work for a single call, but as the azimuth of the source changes over time the inversion procedure would have to be repeated more often to account for apparent changes in the lateral hydrophone offsets that would not be captured by measurements of the spatial noise coherence. One solution would be to incorporate inclinometers that measure both array tilt direction as well as magnitude, so that the effects of physical array tilt and clock offset can be distinguished, accommodated, and even exploited to measure source azimuth.

Other questions include what maximum vertical aperture would still permit ambient noise measurements to be used for measuring clock drift, as the spatial coherence magnitude generally is inversely proportional to sensor separation. For example, vertical-array MFP methods have been demonstrated to work in water depths at least 100 m, for signals up to 1 kHz. Can the ambient noise spatial coherence be used at these water depths as well? Is MFP still possible in a strong current, where flow noise and large array inclinations might make synchronization more difficult? What minimum SNR is needed to adequately time-align the data?

What can be concluded at this stage of development, however, is that a new generation of autonomous recorders has made a new type of underwater acoustic array technically feasible, an "insta-array" where acoustic elements can be arranged quickly into arbitrary configurations using conventional fishing and boating gear, and then time-synchronized using passive acoustic methods. The question remaining is not whether such a system is possible, but under what circumstances such a system is practical.

Acknowledgment

The authors would like to thank M. Johnson of Woods Hole Oceanographic Institution (WHOI), Woods Hole, MA, for advice and encouragement on the "insta-array" concept, G. Deane of Scripps Institution of Oceanography (SIO), San Diego, CA, for inviting the first author to participate in the HARC project, J. Dufour of SIO for helping to calibrate the auxiliary sensors on the autonomous recorders, and B. Gisiner and J. Simmen of Office of Naval Research (ONR), Arlington, VA, for providing project support. Two anonymous reviewers provided constructive feedback on the manuscript.

References

[1] A. B. Baggeroer, W. A. Kuperman, and H. Schmidt, "Matched field processing: Source localization in correlated noise as an optimum parameter estimation problem," *J. Acoust. Soc. Amer.*, vol. 83, pp. 571–587, 1988.

[2] N. O. Booth, P. A. Baxley, J. A. Rice, P. W. Schey, W. S. Hodgkiss, G. L. D'Spain, and J. J. Murray, "Source localization with broad-band matched-field processing in shallow water," *IEEE J. Ocean. Eng.*, vol. 21, no. 4, pp. 402–412, Oct. 1996.

[3] A. M. Thode, G. L. D'Spain, and W. A. Kuperman, "Matched-field processing and geoacoustic inversion of blue whale vocalizations," *J. Acoust. Soc. Amer.*, vol. 107, pp. 1286–1300, 2000.

[4] H. P. Bucker, "Use of calculated sound fields and matched field detection to locate sound sources in shallow water," *J. Acoust. Soc. Amer.*, vol. 59, pp. 368–373, 1976.

[5] D. H. Cato, "Sounds of humpback whales migrating along the Australian coastlines," in *Proc. 15th Int. Conf. Acoust.*, 1995, pp. 219–222.

[6] G. W. Hafner, C. L. Hamilton, W. W. Steiner, T. J. Thompson, and H. E. Winn, "Signature information in the song of the humpback whale," *J. Acoust. Soc. Amer.*, vol. 66, pp. 1–6, 1979.

[7] M. J. Noad and D. H. Cato, "Comparison of acoustic and visual surveying of humpback whales off East Australia," *J. Acoust. Soc. Amer.*, vol. 108, p. 2540, 2000.

[8] ——, "Acoustic surveys of humpback whales: Calibration experiments off the east coast of Australia," *J. Acoust. Soc. Amer.*, vol. 112, p. 2398, 2002.

[9] M. J. Noad, D. H. Cato, M. M. Bryden, M.-N. Jenner, and C. S. Jenner, "Cultural revolution in whale songs," *Nature*, vol. 400, p. 537, 2000.

[10] K. Payne and R. Payne, "Large scale changes over 19 years in songs of humpback whales in Bermuda," *A. Tierpsychol.*, vol. 68, pp. 89–114, 1985.

[11] R. Payne, *Among Whales*. New York: Scribner, 1995.

[12] D. H. Cato, "Songs of humpback whales: The Australian perspective," *Mem. Queensland Museum*, vol. 30, pp. 277–290, 1991.

[13] W. C. Burgess, "The bioacoustic probe: A general-purpose acoustic recording tag (A)," *J. Acoust. Soc. Amer.*, vol. 108, p. 2583, 2000.

[14] W. C. Burgess, P. L. Tyack, B. J. L. Boeuf, and D. P. Costa, "A programmable acoustic recording tag and first results from free-ranging northern elephant seals," *Deep-Sea Res. II*, vol. 45, pp. 1327–1351, 1998.

[15] M. Johnson and P. L. Tyack, "A digital acoustic recording tag for measuring the response of wild marine mammals to sound," *IEEE J. Ocean. Eng.*, vol. 28, no. 1, pp. 3–13, Jan. 2003.

[16] S. Dosso and B. Sotirin, "Optimal array element localization," *J. Acoust. Soc. Amer.*, vol. 106, pp. 3445–3459, 1999.

[17] M. D. Collins and W. A. Kuperman, "Focalization – Environmental focusing and source localization," *J. Acoust. Soc. Amer.*, vol. 90, pp. 1410–1422, 1991.

[18] S. E. Dosso, M. J. Wilmut, and A. L. S. Lapinski, "An adaptive-hybrid algorithm for geoacoustic inversion," *IEEE J. Ocean. Eng.*, vol. 26, no. 3, pp. 324–336, Jul. 2001.

[19] M. Fallat and S. Dosso, "Geoacoustic inversion via local, global, and hybrid algorithms," *J. Acoust. Soc. Amer.*, vol. 105, pp. 3219–3230, 1999.

[20] D. F. Gingras and P. Gerstoft, "Inversion for geometric and geoacoustic parameters in shallow water: Experimental results," *J. Acoust. Soc. Amer.*, vol. 97, pp. 3589–3598, 1995.

[21] B. F. Cron and C. H. Sherman, "Spatial correlation functions for various noise models," *J. Acoust. Soc. Amer.*, vol. 34, pp. 1732–1736, 1962.

[22] W. A. Kuperman and F. Ingenito, "Spatial correlation of surface generated noise in a stratified ocean," *J. Acoust. Soc. Amer.*, vol. 67, pp. 1988–1996, 1980.
[23] J. S. Perkins, W. A. Kuperman, F. Ingenito, L. T. Fialkowski, and J. Glattetre, "Modeling ambient noise in three-dimensional ocean environments," *J. Acoust. Soc. Amer.*, vol. 92, pp. 739–752, 1993.
[24] F. B. Jensen, W. A. Kuperman, M. B. Porter, and H. Schmidt, *Computational Ocean Acoustics*. New York: American Institute of Physics, 1994.
[25] G. Deane and M. J. Buckingham, "Vertical coherence of ambient noise in shallow water overlying a fluid seabed," *J. Acoust. Soc. Amer.*, vol. 102, pp. 3413–3424, 1997.
[26] N. M. Carbone, G. Deane, and M. J. Buckingham, "Estimating the compressional and shear wave speeds of a shallow water seabed from the vertical coherence of ambient noise in the water column," *J. Acoust. Soc. Amer.*, vol. 103, pp. 801–813, 1998.
[27] P. Gerstoft, "Inversion of acoustic data using a combination of genetic algorithms and the Gauss-Newton approach," *J. Acoust. Soc. Amer.*, vol. 97, pp. 2181–2190, 1995.
[28] C. E. Lindsay and N. R. Chapman, "Matched field inversion for geoacoustic model parameters using adaptive simulated annealing," *IEEE J. Ocean. Eng.*, vol. 18, no. 3, pp. 224–231, Jul. 1993.
[29] M. Musil, M. J. Wilmut, and N. R. Chapman, "A hybrid simplex genetic algorithm for estimating geoacoustic parameters using matched field inversion," *IEEE J. Ocean. Eng.*, vol. 24, no. 3, pp. 358–369, Jul. 1999.
[30] S. Dosso, G. Brooke, S. Kilistoff, B. Sotirin, V. McDonald, M. Fallat, and N. Collison, "High-precision array element localization for vertical line arrays in the Arctic Ocean," *J. Acoust. Soc. Amer.*, vol. 23, pp. 365–379, 1998.
[31] W. S. Hodgkiss, D. E. Ensberg, G. L. D'Spain, N. O. Booth, and P. W. Schey, "Comparison of acoustic and non-acoustic methods of vertical line array element localization," in *Proc. MTS/IEEE OCEANS*, San Diego, CA, 1995, pp. 1296–1302.
[32] J.-M. Q. D. Tran and W. S. Hodgkiss, "Array surveying using matched-field processing," *J. Acoust. Soc. Amer.*, vol. 94, pp. 2851–2858, 1993.
[33] A. B. Baggeroer, W. A. Kuperman, and P. N. Mikhalevsky, "An overview of matched-field methods in ocean acoustics," *IEEE J. Ocean. Eng.*, vol. 18, no. 4, pp. 401–424, Oct. 1993.
[34] D. J. Battle, P. Gerstoft, W. A. Kuperman, W. S. Hodgkiss, and M. Siderius, "Geoacoustic inversion of tow-ship noise via near-field-matched-field processing," *IEEE J. Ocean. Eng.*, vol. 28, no. 3, pp. 454–467, Jul. 2003.
[35] J. P. Hermand, "Broad-band geoacoustic inversion in shallow water from waveguide impulse response measurements on a single hydrophone: Theory and experimental results," *IEEE J. Ocean. Eng.*, vol. 24, no. 1, pp. 41–66, Jan. 1999.
[36] D. P. Knobles, R. A. Kock, L. A. Thomspon, K. C. Focke, and P. E. Eisman, "Broadband sound propagation in shallow water and geoacoustic inversion," *J. Acous. Soc. Amer.*, vol. 113, pp. 205–222, 2003.
[37] P. Ratilal, P. Gerstoft, Y. K. Pong, and J. T. Goh, "Inversion of pressure data on a vertical array for seafloor geoacoustic properties," *J. Comput. Acoust.*, vol. 6, pp. 269–289, 1998.
[38] A. M. Thode, M. Zanolin, E. Naftali, P. Ratilal, and N. C. Makris, "Necessary conditions for a maximum likelihood estimate to become asympotically unbiased and attain the Cramer-Rao lower bound, Part II: Range and depth localization of a sound source in an ocean waveguide," *J. Acoust. Soc. Amer.*, vol. 112, pp. 1890–1910, 2002.
[39] M. Zanolin, I. Ingram, A. M. Thode, and N. C. Makris, "Asymptotic accuracy of geoacoustic inversions," *J. Acoust. Soc. Amer.*, vol. 116, pp. 2031–2042, 2004.
[40] P. Gerstoft, "SAGA Users Guide 2.0, An Inversion Software Package" SACLANT Undersea Research Centre, La Spezia, Italy, SM-333, 1997 [Online]. Available: www.mpl.ucsd.edu/people/gerstoft
[41] M. I. Skolnik, *Introduction to Radar Systems*. New York: McGraw-Hill, 1962.
[42] V. A. Zakharov, V. A. Lazerev, A. A. Saltykov, A. D. Sokolov, L. I. Tatrinov, and G. A. Sharonov, "Variation of the spatial correlation coefficient of a sound field in a nonuniform time-varying waveguide," *Akusticheskii Zurnal*, vol. 38, pp. 356–359, 1992.
[43] P. Roux and M. Fink, "Green's function estimation using secondary sources in a shallow water environment," *J. Acoust. Soc. Amer.*, vol. 113, pp. 1406–1416, 2003.
[44] C. H. Harrison and D. G. Simons, "Geoacoustic inversion of ambient noise: A simple method," *J. Acoust. Soc. Amer.*, vol. 112, pp. 1377–1389, 2002.
[45] K. G. Sabra, P. Roux, A. M. Thode, G. L. D'Spain, W. S. Hodgkiss, and W. A. Kuperman, "Using ocean ambient noise for array self-localization and self-synchronization," *IEEE J. Ocean. Eng.*, 2006, to be published.
[46] P. Roux, W. A. Kuperman, and N. group, "Extracting coherent wave fronts from acoustic ambient noise in the ocean," *J. Acous. Soc. Amer.*, vol. 116, pp. 1995–2003, 2004.
[47] C. H. Harrison, "Sub-bottom profiling using ocean ambient noise," *J. Acous. Soc. Amer.*, vol. 115, pp. 1505–1515, 2004.
[48] K. G. Sabra, P. Gerstoft, P. Roux, W. A. Kuperman, and M. C. Fehler, "Extracting time-domain Greens function estimates from ambient seismic noise," *AGU Geophys. Res. Lett.*, vol. 32, p. L03310, 2005.
[49] J. H. Goodman, *Statistical Optics*. New York: Wiley, 1985.
[50] W. A. Watkins and W. E. Schevill, "Sound source location by arrival times on a non-rigid three-dimensional hydrophone array," *Deep-Sea Res.*, vol. 19, pp. 691–706, 1972.
[51] M. J. Noad and D. H. Cato, "A combined acoustic and visual survey of humpback whales off Southeast Queensland," *Mem. Queensland Museum*, vol. 47, pp. 507–523, 2001.
[52] R. Payne and S. McVay, "Songs of humpback whales," *Science*, vol. 173, pp. 585–597, 1971.
[53] F. L. Macknight, D. H. Cato, M. J. Noad, and G. C. Grigg, "Qualitative and quantitative analyses of the song of the East Australian population of humpback whales," *Mem. Queensland Museum*, vol. 47, pp. 525–537, 2002.
[54] T. F. Norris, M. Mc Donald, and J. Barlow, "Acoustic detections of singing humpback whales (Megaptera novaeangliae) in the Eastern North Pacific during their northbound migration," *J. Acoust. Soc. Amer.*, vol. 106, pp. 506–514, 1999.
[55] T. C. Yang, "Environmental effects on phase coherent underwater acoustic communications: A perspective from several experimental measurements," in *Proc. Conf. High Frequency Ocean Acoust.*, La Jolla, CA, 2004, pp. 90–97.
[56] J.-M. Q. D. Tran and W. S. Hodgkiss, "Sound-speed profile inversion using a large aperture vertical line array," *J. Acoust. Soc. Amer.*, vol. 93, pp. 803–812, 1993.
[57] G. L. D'Spain, J. J. Murray, W. S. Hodgkiss, N. O. Booth, and P. W. Schey, "Mirages in shallow water matched field processing," *J. Acoust. Soc. Amer.*, vol. 105, pp. 3245–3265, 1999.
[58] G. A. Grachev, "Theory of acoustic field invariants in layered waveguide," *Acoust. Physics*, vol. 39, pp. 33–35, 1994.

Aaron M. Thode (M'01) received the B.S. degree in physics and M.S. degree in electrical engineering from Stanford University, Stanford, CA, in 1993 and the Ph.D. degree in oceanography from the Scripps Institution of Oceanography, University of California, San Diego, in 1999.

He was a Postdoctoral Scholar in ocean engineering at the Massachusetts Institute of Technology (MIT), Cambridge, from 1999 to 2001. He is currently an Associate Research Scientist at the Marine Physical Laboratory at Scripps Institution of Oceanography.

Dr. Thode received the Office of Naval Research Acoustic Entry-Level Faculty Award, in 2002, and the A. B. Wood Medal from the U.K. Institute of Acoustics/Acoustical Society of America, in 2005. He is a member of the Acoustical Society of America, the American Geophysical Union, and the Society for Marine Mammalogy.

Peter Gerstoft received the M.Sc. and the Ph.D. degrees in civil engineering from the Technical University of Denmark, Lyngby, Denmark, in 1983 and 1986, respectively, and the M.Sc. degree from the University of Western Ontario, London, ON, Canada, in 1984.

From 1987 to 1992, he was with Ödegaard and Danneskiold-Samsöe, Copenhagen, Denmark, working on forward modeling and inversion for seismic exploration, with a year as a Visiting Scientist at the Massachusetts Institute of Technology (MIT), Cambridge. From 1992 to 1997, he was a Senior Scientist at SACLANT

Undersea Research Centre, La Spezia, Italy, where he developed the SAGA inversion code, which is used for ocean acoustic and electromagnetic signals. Since 1997, he has been with the Marine Physical Laboratory, University of California, San Diego. His research interests include global optimization, modeling and inversion of acoustic, and elastic and electromagnetic signals.

Dr. Gerstoft is a Fellow of the Acoustical Society of America and elected member of the International Union of Radio Science, Commission F.

William C. Burgess (M'93) received the M.S. degree in electrical engineering with a focus on computer operating systems, in 1987 and the Ph.D. degree in electrical engineering for studies of the effects of precipitating radiation-belt electrons on subionospheric radio propagation, in 1993, both from Stanford University, Stanford, CA.

He moved to ocean science and acoustics under a Postdoctoral Scholarship at the Woods Hole Oceanographic Institution, Woods Hole, MA, and a Postdoctoral Fellowship at the Monterey Bay Aquarium Research Institute, University of California, San Diego. During the latter appointment, he designed and built the first microprocessor-based acoustic recording tag for attachment to marine wildlife, applying it with northern elephant seals on the California coast. Since joining Greeneridge Sciences, Goleta, CA, as a Senior Research Engineer in 1998, he has focused on the understanding and mitigation of human acoustic impacts on marine and riverine species. With support from the Office of Naval Research he developed the bioacoustic probe acoustic recorders used in this study, and is presently designing the next generation of this instrument.

Dr. Burgess is a member of the Acoustical Society of America, the American Geophysical Union, and the Society for Marine Mammalogy.

Karim G. Sabra received the M.S. and Ph.D. degrees in mechanical engineering from the Mechanical Engineering Department, University of Michigan, Ann Arbor, in 2000 and 2003, respectively. His thesis research investigated the broadband performance of an acoustic time-reversing array retrofocusing in a shallow-water sound channel, as well as application of time-reversal signal processing to the blind deconvolution problem.

He was a Postdoctoral Researcher at the Marine Physical Laboratory, University of California, San Diego, from 2003 to 2005, working on noise-based tomography using oceans microseisms and ambient noise and iterative time-reversal methods for target detection. Currently, he is a Project Scientist at the Marine Physical Laboratory. His research interests include time-reversal acoustics, geophysics, coherent processing of ambient noise, and buried target detection.

Melania Guerra received the B.S. degree in mechanical engineering from the Universidad de Costa Rica, Costa Rica, in 2001 and is currently working towards the M.S. degree at the Marine Physical Laboratory, Scripps Institution of Oceanography, San Diego, CA.

In 2002, she was an Intern at the Advanced Space Propulsion Lab, NASA Johnson Space Center, Houston, TX, performing experiments on the high-temperature superconducting magnets (HTSM) of the VASIMR rocket.

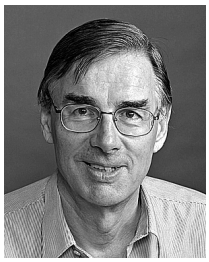
M. Dale Stokes received the Ph.D. degree in oceanography from the Scripps Institution of Oceanography, University of California, San Diego, in 1995.

His major fields of research include air-sea gas exchange and shallow-water physical oceanography. He was a Principal Organizer of the Humpback Acoustic Research Consortium (HARC).

Michael Noad received the degree in veterinary science from the University of Queensland, St. Lucia, Qld., Australia, in 1990 and the Ph.D. degree on humpback whale song from the Faculty of Veterinary Science, University of Sydney, Sydney, N.S.W., Australia, in 2002, supervised by Prof. M. Bryden and Dr. D. Cato. His thesis involved acoustic and behavioral studies of the whales migrating along the East Coast of Australia. He documented when the song of east Australian population was rapidly replaced by that of the west Australian population, a cultural change the scale and nature of which had not been previously recorded in any species of nonhuman animal.

He was a Postdoctoral Fellow in the School of Life Science, University of Queensland and became a Principal Organizer of the Humpback whale Acoustic Research Collaboration (HARC). Other research interests include determining the relative and absolute abundance of the east Australian humpback population, using acoustics as a survey tool for abundance estimates, and the anatomy and pathology of cetacean ears.

Dr. Noad is also a member of the South Pacific Whale Research Consortium.



Douglas H. Cato received the Ph.D. degree in electrical engineering from the University of Sydney, Sydney, Australia.

Currently, he is a Principal Research Scientist at the Defence Science and Technology Organization, Sydney, Australia, and has an adjunct appointment at the University of Sydney Institute of Marine Science, Sydney, Australia. His research interests include ambient noise, marine bioacoustics, and effects of noise on marine animals.

Dr. Cato is a Fellow of the Acoustical Society of America.