Ranging bowhead whale calls in a shallow-water dispersive waveguide

Shima H. Abadi[a)]
Department of Mechanical Engineering, University of Michigan, Ann Arbor, Michigan 48109

Aaron M. Thode
Marine Physical Laboratory, Scripps Institution of Oceanography, University of California at San Diego, La Jolla, California 92093

Susanna B. Blackwell
Greeneridge Sciences, Inc., 6160-C Wallace Becknell Road, Santa Barbara, California 93117

David R. Dowling
Department of Mechanical Engineering, University of Michigan, Ann Arbor, Michigan 48109

(Received 25 November 2013; revised 27 March 2014; accepted 16 May 2014)

This paper presents the performance of three methods for estimating the range of broadband (50–500 Hz) bowhead whale calls in a nominally 55-m-deep waveguide: Conventional mode filtering (CMF), synthetic time reversal (STR), and triangulation. The first two methods use a linear vertical array to exploit dispersive propagation effects in the underwater sound channel. The triangulation technique used here, while requiring no knowledge about the propagation environment, relies on a distributed array of directional autonomous seafloor acoustics recorders (DASARs) arranged in triangular grid with 7 km spacing. This study uses simulations and acoustic data collected in 2010 from coastal waters near Kaktovik, Alaska. At that time, a 12-element vertical array, spanning the bottom 63% of the water column, was deployed alongside a distributed array of seven DASARs. The estimated call location-to-array ranges determined from CMF and STR are compared with DASAR triangulation results for 19 whale calls. The vertical-array ranging results are generally within ±10% of the DASAR results with the STR results providing slightly better agreement. The results also indicate that the vertical array can range calls over larger ranges and with greater precision than the particular distributed array discussed here, whenever the call locations are beyond the distributed array boundaries. © *2014 Acoustical Society of America*.
[http://dx.doi.org/10.1121/1.4881924]

PACS number(s): 43.30.Sf, 43.30.Bp, 43.60.Jn [SED] Pages: 130–144

I. INTRODUCTION

The localization of marine mammals is important for biological studies and for assessments of the impact of anthropogenic activities on the marine environment. For example, by plotting the spatial distribution of vocalizing animals around an anthropogenic activity, one can infer at what received levels avoidance behavior is triggered. A specific case of concern is the potential behavioral responses of migrating bowhead whales to various kinds of oil and gas industrial activity in the Arctic Ocean, particularly seismic airgun surveys (Richardson *et al.*, 1986; Blackwell *et al.*, 2013).

Passive acoustic monitoring (PAM) has become an increasingly popular method for localizing certain marine mammal species that are acoustically active. The bowhead whale (*Balaena mysticetus*) is one such species, generating a large variety of calls, including impulsive and frequency-modulated (FM) sweeps, during its spring and fall migrations between the Bering and Beaufort Seas (Clark and Johnson, 1984; Cummings and Holliday, 1987; Moore *et al.*, 2006; Thode *et al.*, 2012). Traditional PAM methods use differences in arrival times of a signal between receivers to estimate a position (for recent work, see Nosal, 2013; von Benda-Beckmann *et al.*, 2013); however, to localize and track positions over a significant area, these methods generally require the deployment of large numbers of widely distributed sensors, which must then be time-synchronized. For example, the tracking of bowhead whales in the Arctic Ocean, summarized in Clark and Ellison (2000), used arrival-time differences of transient bowhead sounds detected on a sparse sonobuoy array. More recent efforts include the use of ocean bottom seismometer networks for tracking fin and blue whales in the northeast Pacific Ocean (Wilcock, 2012) and the development of "M3R" software, used by U.S. Navy test ranges to perform two-dimensional (2D) and (3D) hyperbolic localization (Morrissey *et al.*, 2006).

An alternative, less common method, is to use either short-baseline arrays or acoustic vector sensors to measure the azimuth of sounds from several locations and then to use triangulation to position the animal. The directional autonomous seafloor acoustic recorders (DASARs) currently used by Greeneridge Sciences Inc. (Greene *et al.*, 2004) apply this technique for localizing bowhead whale calls in the shallow

[a)]Author to whom correspondence should be addressed. Electronic mail: shimah@umich.edu

 0001-4966/2014/136(1)/130/15/$30.00

waters of the Beaufort Sea. While this technique eliminates the need to precisely time-synchronize the instruments, it still requires the deployment of numerous instrument packages across a wide region and also requires considerable field and analysis time to determine the orientation of the sensors on the ocean floor.

More recent work has explored how to exploit acoustic multipath arrivals to provide additional localization information. In shallow water, Bayesian statistics and autocorrelation methods involving multipath ray acoustics have been used to track walruses (Rideout *et al.*, 2013) and baleen whales (Valtierra *et al.*, 2013). In deep water, surface-reflected and/or refracted ray paths have been used to track sperm and beaked whales using fewer hydrophones than standard techniques (Tiemann *et al.*, 2004; Tiemann *et al.*, 2006; Nosal and Frazer, 2007; Mathias *et al.*, 2013; Zimmer, 2013). An extreme case of multipath propagation arises when relatively low-frequency sounds propagate in shallow water environments. Here, no direct path arrival exists, and the propagating signal is appropriately modeled as a set of normal modes. Each mode has a different effective group velocity, so, in principle, if the different modal arrivals can be separated in time, the range of the call can be determined using measurements at a single point (D'Spain *et al.*, 1995; Wiggins *et al.*, 2004; Munger *et al.*, 2011). Unfortunately, this approach is manually intensive and only works when the source is a sufficient distance from the sensor to permit temporal isolation of modal arrivals, and when sufficient environmental information is available to permit computation of the frequency-dependent mode group velocities.

The use of a vertical array allows isolation of the various modal arrivals even when they overlap in time; this in turn can be used to infer the range and depth of an acoustic source from a single deployment. Furthermore, the required environmental information can often be deduced from the acoustic signal itself via "geoacoustic inversion" (Gerstoft, 1994; Gerstoft and Gingras, 1996; Ratilal *et al.*, 1998; Thode *et al.*, 2000; Thode *et al.*, 2006). An enormous literature exists on "matched-field processing" (MFP), a technique that uses extensive propagation modeling to localize compact acoustic sources on passive acoustic arrays (e.g., Baggeroer *et al.*, 1993). While the technique works for narrowband sounds, decades of experience have found that considerable information about the propagation environment is necessary for robust processing, particularly at lower signal-to-noise ratios (SNRs) (Bucker, 1976; see Jensen *et al.*, 1994). Direct measurements of this information are unlikely to be available (Yoo and Yang, 1999; Zurk and Ward, 2003). For broadband sounds typical of marine mammals, this environmental information can be deduced from the acoustic signals themselves, and thus combined MFP-plus-geoacoustic inversion methods have been successfully applied to blue and humpback whales (Thode *et al.*, 2000; Thode *et al.*, 2006).

If the vertical array spans a sufficient fraction of the water column, and if the signals are of sufficient bandwidth, more computationally efficient mode filtering methods can be applied to a vertical array (Shang *et al.*, 1985; Yang, 1987; Lu *et al.*, 1993). This technique converts acoustic measurements into the amplitudes and phases of the excited propagating modes (Lo *et al.*, 1983; Buck *et al.*, 1998; Shang *et al.*, 1985). A broadband remote source's location can then be estimated by matching these amplitudes and phases with those obtained from modeling propagating modes, a technique known as matched-mode processing (Yang, 1989; Krolik, 1992; Collison and Dosso, 2000). The concept has also been called adaptive modal back-propagation and has recently been applied to tracking sei whales along the U.S. East Coast (Lin *et al.*, 2012). A major challenge of any mode filtering technique is that the vertical array must span sufficient aperture in the water column to exploit the orthogonal relationships between the modes. In addition, like all the other shallow-water dispersive tracking methods, sufficient knowledge about the environment is required to apply the mode-filtered results toward matched-mode processing or back-propagation.

This paper uses both simulations and Arctic data to compare the performance of three PAM ranging methods when applied to bowhead whale calls in the 50–500 Hz range, propagating in a nominally 55-m-deep sound channel. Two techniques require the use of a vertical array: Conventional mode-filtering (CMF) and synthetic time reversal (STR), a blind deconvolution technique (Sabra and Dowling, 2004; Sabra *et al.*, 2010; Abadi *et al.*, 2012). The vertical array results are benchmarked against those obtained from standard triangulation using separate widely distributed bottom-mounted DASARs (Greene *et al.*, 2004; Thode *et al.*, 2012). The vertical array methods offer potential advantages over sparse DASAR arrays (and other traditional time-of-arrival techniques) in that they could provide more precise range estimates over a larger area, using a single deployment (vs seven DASAR deployments). However, the vertical array methods require more detailed knowledge of the propagation environment; this makes them vulnerable to range bias, and they must compensate for the fact that a practical vertical array deployment in the Arctic cannot span the full water column because of the potential problems caused by pack ice. Therefore a technique like the "array invariant" (Lee and Makris, 2006), which requires precise bearing measurements of propagating energy, is precluded.

CMF and STR are similar but not identical. Both techniques involve maximizing the cross-correlation of mode-filtered signals to determine source-array range. Both are suitable for use in underwater sound channels with modal dispersion in the bandwidth of interest when the environment is not completely known. Both require a modest amount of environmental information at the receiving array. And both perform best in range-independent environments. However, there is one important difference between the two techniques. The CMF-based technique relies on mode-shape weighting of the array measurements to estimate signal amplitude *and* phase in the signal bandwidth, while the STR-based technique relies on mode-shape weighting to determine only signal phase in the signal bandwidth. Furthermore the STR method exploits all modal arrivals when determining the signal, whereas CMF only uses the signal contained in a single mode. Hence STR is expected to be more robust for low SNR signals conveyed by several modes.

The remainder of this paper is divided into four sections. The next section presents the background and mathematical

formulation of CMF, mode-based STR, and DASAR ranging techniques in a dispersive environment. Section III presents CMF and STR results from simulated acoustic propagation in a simple environment that mimics the Arctic Ocean environment discussed in Sec. IV, which presents results obtained from bowhead whale calls recorded in 2010 using a 12-element vertical array deployed in the bottom 63% of a 55-m-deep water column. Section IV also compares these results with those from DASAR ranging methods. The last section evaluates the relative performance of all three techniques.

II. BACKGROUND ON THREE RANGING METHODS

This section presents the mathematical formulation of three different whale call ranging techniques. The first technique is based on mode filtering, which has a long history (Yang, 1987; Buck *et al.*, 1998). The second technique is an extension of mode-based STR (Sabra and Dowling, 2004), which can be used for sound source localization when the medium is not dispersive and basic environmental information is available at the receiving array (Abadi *et al.*, 2012). Here an extension of mode-based STR is used for ranging whale calls in a dispersive medium.

The third technique is whale-call triangulation using DASAR recordings deployed as a sparse array in a triangular grid (Greene *et al.*, 2004; Thode *et al.*, 2012). DASAR triangulation has an established performance record, but this performance deteriorates rapidly with range for calls detected outside the boundary of the sparse array.

Figure 1 illustrates the vertical array geometry and defines some of the notation used in the following text. The array geometry displayed here is derived from the actual deployment geometry described in Sec. IV B, and the Pekeris waveguide environment displayed here is obtained from a geoacoustic inversion procedure described in Sec. IV C.

A. CMF

There are a variety of algorithms for implementing conventional mode-filtering/matched-mode processing; here we introduce a format that permits easy comparison with the STR method. In the current study, CMF was used to estimate the source signal $\hat{S}_m(\omega)$ in the frequency domain from the

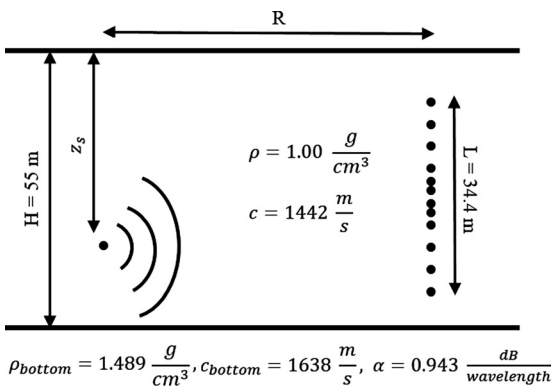

FIG. 1. Array geometry and range-independent Pekeris waveguide model used in simulations and experimental data analysis. The 12-element receiving array is centered at a depth of 28.8 m in 55 m water depth. The environment shown was inverted from a 7.5 km range whale call listed in Table I.

Fourier transforms $P_j(\omega)$ of signals $p_j(t)$ recorded by the N elements of a vertical receiving array, where j is the array hydrophone index $(1 \le j \le N)$. Here the array-element weights were determined from the vertical shape of the mth propagating acoustic mode

$$\hat{S}_m(\omega) = \sum_{j=1}^{N} W_{j,m} P_j(\omega), \tag{1}$$

$$\text{where } W_{j,m} = \psi_m(z_j), \tag{2}$$

ψ_m is the mth propagating mode function, which can be either approximated from an ideal waveguide model or derived from the data using geoacoustic inversion methods as described in Sec. IV C. This choice of array-element weights, $W_{j,m}$, may be able to isolate the sound that propagates via a single mode if both the vertical array aperture and spatial sampling are sufficient to exploit the orthogonal relationship between modes,

$$\sum_{j=1}^{N} (z_{j+1} - z_j)\psi_n(z_j)\psi_m(z_j) \sim \delta_{mn}. \tag{3}$$

More sophisticated weighting schemes are possible for the same task (Buck *et al.*, 1998). However, for the current investigation, results from such schemes were largely the same as those from the elementary mode-filtering approach, so Eqs. (1) and (2) are used here for simplicity.

Equation (1) provides an estimate of the Fourier transform of the source signal. Because $P_j(\omega)$ is really a convolution of the sound channel's Green's functions and the actual source signal, Eq. (1) can be re-written as follows:

$$\hat{S}_m(\omega) = |\hat{S}_m(\omega)| e^{i\{\hat{\varphi}_s(\omega) + \Gamma_m(\omega)\}}, \tag{4}$$

where $\hat{\varphi}_s$ is the original signal's source phase, and $\Gamma_m(\omega)$ is an extra phase arising from the source-to-array propagation via the mth mode. If Eq. (3) holds, then this extra phase is well approximated by

$$\Gamma_m(\omega) = \text{Re}\{k_m\}R, \tag{5}$$

where $k_m = \omega/c_{p,m}$ is the horizontal wave number of the mth mode, $c_{p,m}$ is the phase speed of mth mode, and R is the horizontal range from the remote source to the array.

If the cutoff frequency of the mth mode is well below the lowest signal frequency, the phase speed of the mth mode can be assumed independent of frequency, and the extra phase given by Eq. (5) will be a linear function of frequency for a fixed source to receiver geometry. This linear-in-frequency phase is equivalent to a time shift in the time domain, and thus the estimated source signal $\hat{S}_m(\omega)$ will simply be a time-shifted version of the original source signal, excluding attenuation effects: a so-called "blind deconvolution" (Sabra and Dowling, 2004).

However, the frequency range of most bowhead whale calls extends down below 150 Hz, a highly dispersive regime for the first four propagating modes, as illustrated in Fig. 2, which shows phase speed results from the propagation model

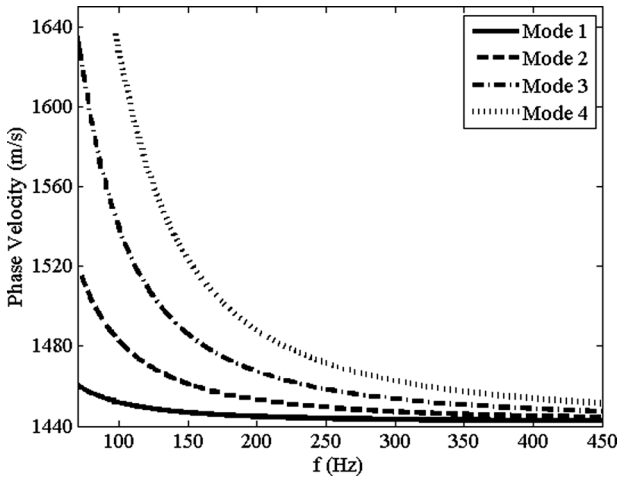

FIG. 2. Phase velocities of the first four modes vs frequency calculated by KRAKEN for the environment shown in Fig. 1. It shows that all modes, including mode 1, are dispersive below 150 Hz.

KRAKEN (Porter and Reiss, 1984) for the environment shown in Fig. 1. Under these circumstances, the phase speed of the mth mode becomes a function of frequency, so $\hat{S}_m(\omega)$ is a phase-distorted version of the original signal. Thus STR blind deconvolution is not possible over this frequency range, but an estimate of the source range becomes feasible if the signal contains two or more propagating modes.

The range estimation procedure involves applying Eq. (1) twice to a given received signal using two different modes, m and n, to produce two source signature estimates, $\hat{S}_m(\omega)$ and $\hat{S}_n(\omega)$. The phase of their frequency-domain cross-correlation is then of interest

$$\arg\left\{\hat{S}_m(\omega)\hat{S}_n^*(\omega)\right\} = \left\{\hat{\varphi}_{s,m}(\omega) - \hat{\varphi}_{s,n}(\omega)\right\} + \left\{k_m(\omega) - k_n(\omega)\right\}R. \quad (6)$$

If the medium were non-dispersive, and both modes were excited by the same sound source, $\hat{S}_m(\omega)$ and $\hat{S}_n(\omega)$ should be identical, and the signal phase estimates in Eq. (6) should cancel. However, $k_m(\omega) - k_n(\omega) \equiv k_{mn}(\omega)$ is a function of frequency in dispersive environments, and such dispersion decreases the peak cross-correlation between $\hat{S}_m(\omega)$ and $\hat{S}_n(\omega)$. If $\{k_m(\omega) - k_n(\omega)\}R \equiv k_{mn}R$, the dispersive portion of the phase, can be estimated and removed from Eq. (6), then the m- and n-mode-reconstructed source signals should become highly correlated. If some environmental knowledge is available that permits the difference between horizontal wavenumbers to be estimated, then the range R can be obtained from $k_{mn}(\omega)R$.

Grachev (1993) found that the difference between horizontal wavenumbers in a waveguide was proportional to $\omega^{-1/\beta}$, where β, the "waveguide invariant," is $\sim$1 for waveguides defined by reflecting boundaries (i.e., most shallow water waveguides). So the second dispersive term in Eq. (6) can simply be modeled as

$$k_{mn}(\omega)R \approx 2\pi\Delta_{mn}/\omega, \quad (7)$$

where Δ_{mn} is a fitting constant with units of s^{-1}. Here Δ_{mn} can be determined as the value that maximizes the cross-correlation between $\hat{S}_m(\omega)$ and $\hat{S}_n(\omega)$ when the phase given by Eq. (6) is adjusted by Eq. (7)

$$\hat{S}_m(\omega)\hat{S}_n^*(\omega)\exp\{-i2\pi\Delta_{mn}/\omega\} = |\hat{S}_m(\omega)||\hat{S}_n(\omega)|\exp\{i\left[\hat{\varphi}_{s,m}(\omega) - \hat{\varphi}_{s,n}(\omega)\right] + ik_{mn}(\omega)R - i2\pi\Delta_{mn}/\omega\}. \quad (8)$$

Taking the inverse Fourier transform of Eq. (8) yields a dispersion-adjusted time-domain cross-correlation. When Δ_{mn} is adjusted to maximize this cross-correlation, the resulting value of Δ_{mn} nearly eliminates $k_{mn}(\omega)R$ from Eq. (8). Thus the range of the calling whale may then be found by minimizing the difference or error between $2\pi\Delta_{mn}/\omega$ and $k_{mn}(\omega)R_e$. If the propagation model is accurate enough, the value of the following mean-square error (MSE) should be minimum when R_e is the true range

$$\mathrm{MSE}(R_e) = \sum_{\omega=\omega_1}^{\omega_2}\left[k_{mn}(\omega)R_e - \frac{2\pi\Delta_{mn}}{\omega}\right]^2. \quad (9)$$

Equations (8) and (9) together are reminiscent of the equations for "matched-mode processing" (MMP) [e.g., Eq. (4) in Shang (1989)]. However, in the present case, the range is estimated using only two modes, whereas in standard MMP, Eq. (9) would effectively be summed over all mode combinations as well as frequency, with each mode contribution weighted by the relative mode excitations detected on the vertical array. In this paper, Eq. (9) will only be applied to modes 1 and 2 ($m=1$ and $n=2$) or modes 2 and 3 ($m=2$ and $n=3$). The motivation behind this choice is that these modes are the most robust to mode filtering [e.g., they best fit the criteria of Eq. (3) for an incomplete array] and to uncertainties in the bottom composition (because, for a fixed acoustic frequency, the lowest order modes have the least interaction with the ocean bottom). Thus restricting Eq. (9) to the first three modes enhances range estimation robustness at the cost of sacrificing potential depth estimation resolution.

Equations (8) and (9) are based on three significant assumptions about the waveguide. First, Eqs. (8) and (9) assume that the difference between $k_{mn}(\omega)R$ and $k_{mn}(\omega+\Delta\omega)R$ (adjacent frequency bins of the FFT of the signal) is less than 2π. Second, the equations assume a range-independent environment, particularly a constant ocean depth. As will be seen in Sec. IV, the ocean bathymetry in the Arctic is effectively flat over the ranges in question out to at least 25 km range. Third, the equations assume the true depth of the sound channel is known. In particular, Grachev (1993) notes that $\Delta_{mn} \propto D^{-2}$, where D is the waveguide depth. Thus if the estimated ocean depth D_{model} is not correct, Eq. (9) will still converge to the same MSE value but at a different $R_e = R_{true}(D_{true}/D_{model})^2$. Thus any range estimation scheme based on modal dispersion, regardless of whether a vertical array is involved or not, will be vulnerable to bias arising from incomplete knowledge of the bottom bathymetry.

B. STR

STR is a simple array signal-processing technique for simultaneously estimating the waveform of the unknown

source signal and the unknown source-to-array transfer functions in an unknown multipath environment. Its formulation appears in Sabra and Dowling (2004), so only final formulae are provided here. Mode-based STR generates $\hat{S}_m(\omega)$ from

$$\hat{S}_m(\omega) = \sum_{j=1}^{N} \left\{ \tilde{P}_j(\omega) e^{-i\alpha_m(\omega)} \right\}^* P_j(\omega), \tag{10}$$

where $\tilde{P}_j(\omega)$ is the normalized received signal, $\tilde{P}_j(\omega) = P_j(\omega)/\sqrt{\sum_{j=1}^{N} |P_j(\omega)|^2}$, and $\alpha_m(\omega)$ is a phase correction constructed from Eqs. (1) and (2)

$$\alpha_m(\omega) = \arg\left\{ \sum_{j=1}^{N} \psi_m(z_j) P_j(\omega) \right\} \approx \hat{\varphi}_s(\omega) + k_m R. \tag{11}$$

Unlike the CMF approach, STR-based signal reconstruction—as represented by Eqs. (10) and (11)—does not rely on Eqs. (1) and (2) to determine the signal amplitude.

After the source signal is reconstructed from two different modes using Eqs. (10) and (11), the remaining steps for whale-call ranging are exactly the same as for the CMF approach. Thus the only difference between CMF- and STR-based whale call ranging is that the estimated Fourier transform of the signal given by Eq. (1) is replaced by Eq. (10) for STR. As discussed in Sabra and Dowling (2004), Eq. (10) uses all modal arrivals to estimate the spectrum, $|\hat{S}_m(\omega)|$, of the source signal, whereas mode-filtering only uses the signal energy of mode m. Thus under low SNR conditions, one might expect the $|\hat{S}_m(\omega)||\hat{S}_n(\omega)|$ terms in the cross-correlation of Eq. (8) to be reconstructed more accurately using STR, and thus the ranging performance of STR might be better under noisy ocean conditions.

C. DASARs

A DASAR instrument package records three channels of acoustic data: One omnidirectional and two directional. It estimates the bearing of a transient acoustic signal consisting of L samples by taking the arctangent of the ratio of the mean active acoustic intensities obtained from the three channels (Greene *et al.*, 2004)

$$\phi_{true} = \phi_{cal} + \tan^{-1}\left[I_y/I_x\right], \tag{12}$$

where the arctangent is defined over four quadrants, and ϕ_{cal} is a "calibration" bearing that relates the orientation of the DASAR directional axes to true north. The mean active acoustic intensity, I_k, is defined as

$$I_k = \frac{1}{L} \sum_{l=1}^{L} p(t_l) v_k(t_l) \tag{13}$$

with $p(t_l)$ being the time series from the pass-band-filtered omnidirectional channel and $v_k(t_l)$ being the filtered time series from one of the directional channels.

The bearings from multiple DASARs are used to estimate a robust maximum-likelihood position of the calling animal (Lenth, 1981a; Greene *et al.*, 2004). A Huber weighting function (Huber, 1964) was used to suppress directional outliers, using a tuning parameter of 1.5 and incorporating 100 bootstrapped estimates of the concentration parameter κ (Lenth, 1981b). The final output is a location bounded by a 90% confidence ellipse. Further details on how the bearings from individual DASARs are combined to generate a position and position uncertainty ellipse are provided in Greene *et al.* (2004).

III. SIMULATIONS OF VERTICAL ARRAY RANGING PERFORMANCE

To understand and predict the performance of CMF and STR in ranging bowhead whale calls, signals that mimic these calls were constructed and propagated through the environment shown in Fig. 1 to simulate the signals received by the vertical array deployed in the Arctic environment. A 12-element vertical array with non-uniform spacing was modeled with the shallowest element being at 14.5 m and the deepest being at 48.9 m, thus spanning 63% of the water column. The element spacing was not uniform; it instead followed the actual array geometry discussed in Sec. IV A.

For the simulations, a linear chirp signal with 60–300 Hz bandwidth was broadcast at 10 km range from a source depth of 45 m. Figure 2 shows the phase speeds for the first four modes calculated for the simple two-layer Pekeris waveguide shown in Fig. 1. All are dispersive over this signal bandwidth. Even the first propagating mode is dispersive below 150 Hz. The shape functions for the first three Pekeris modes are plotted in Fig. 3 for 300 Hz with symbols placed at the depths of the array elements.

Sample CMF and STR comparison results for (1,2) and (2,3)-mode weighting pairs are shown in Fig. 4 with thick and thin lines, respectively. Here white noise has been added to the simulation to generate a 10.5 dB SNR on an individual element, and the array has been modeled with a 2° vertical inclination to simulate a more realistic experimental situation. Figure 4(a) shows the cross-correlation coefficients between $\hat{S}_1(\omega)$ and $\hat{S}_2(\omega)e^{i2\pi\Delta_{12}/\omega}$, and $\hat{S}_2(\omega)$ and

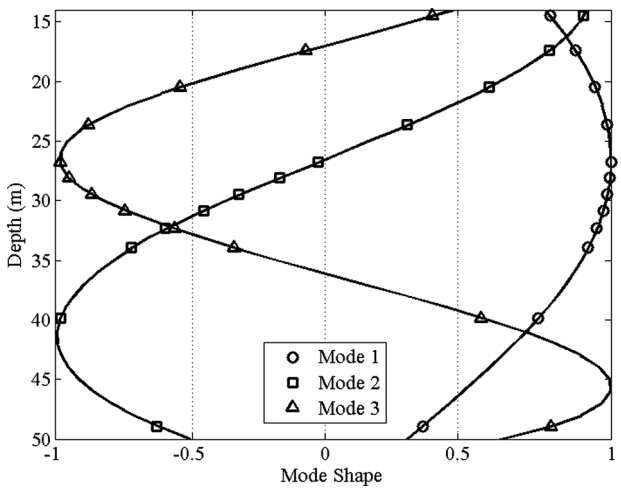

FIG. 3. Normalized mode shapes of the first three modes at receivers' depths at 300 Hz. Mode 2 has a null at 26.5 m depth, and mode 3 has nulls at 17 and 36 m depth.

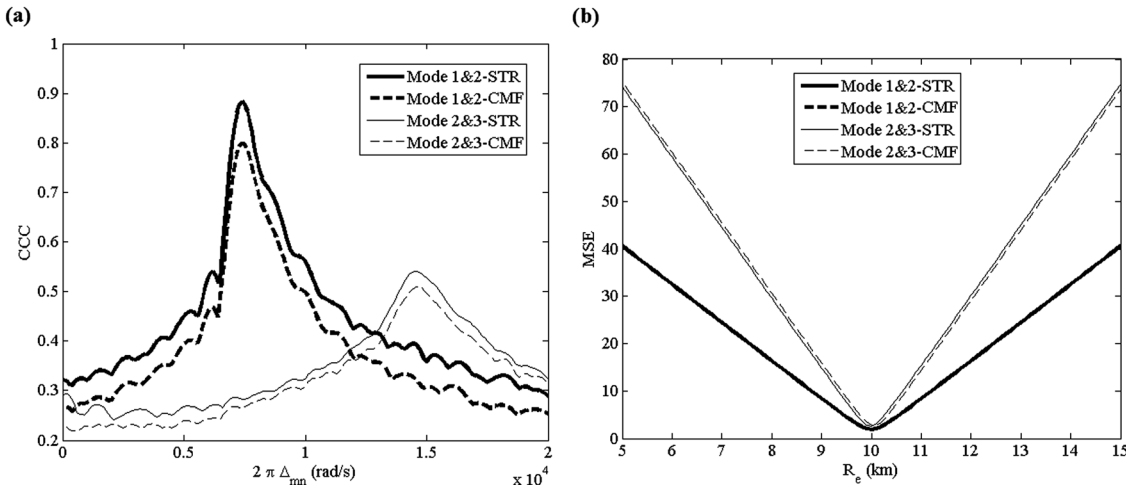

FIG. 4. (a) Simulated cross-correlation coefficient between $\hat{S}_1(\omega)$ and $\hat{S}_2(\omega)e^{i2\pi\Delta_{12}/\omega}$ vs Δ_{12} (thick lines) and $\hat{S}_2(\omega)$ and $\hat{S}_3(\omega)e^{i2\pi\Delta_{23}/\omega}$ vs Δ_{23} (thin lines). (b) Error calculated from Eq. (9) vs estimated range using the (1,2) mode pair and (2,3) mode pair (thick and thin lines, respectively). The respective STR and CMF results are indicated by solid and dashed lines. STR has a slightly higher cross-correlation coefficient than CMF. The simulated source signal is a 60–300 Hz chirp at 45 m depth and 10 km source-to-array range. The vertical array inclination mismatch is 2° for this and the following simulations.

$\hat{S}_3(\omega)e^{i2\pi\Delta_{23}/\omega}$ vs Δ_{12} and Δ_{23} for the CMF and STR ranging techniques, where the mode weights applied in Eq. (2) assume no inclination. For both techniques, there is a maximum in the cross-correlation coefficient (CCC) near $\Delta_{12} \approx 7000$ rad/s and $\Delta_{23} \approx 14\,500$ rad/s. Figure 4(b) shows the MSE from Eq. (9) evaluated at these values of Δ_{12} and Δ_{23}. It shows that the (1,2)-mode pair produces a minimum at $R_e = 10$ km, the correct simulated source-array range, and (2,3)-mode pair produces a minimum at 10.1 and 10 km for CMF and STR, respectively.

Figure 4 shows that mode pairs other than modes 1 and 2 (like modes 2 and 3) can be successfully used for CMF and STR ranging when the mode pair is orthogonal across the array aperture, and both modes are excited by the sound source. For the 12-element vertical array simulated here, the value calculated at 80 Hz by Eq. (3) between mode 1 and 2 is −0.08, between mode 2 and 3 is 0.008, and between mode 1 and 3 is −0.28. So the first and second modes and the second and third modes are approximately orthogonal across the array aperture. However, the first and third modes are not, so ranging success in this investigation was only possible when mode 2 was adequately excited. A sound source at a depth of 17 or 36 m does not excite mode 3, and a sound source at a depth of 26.5 m does not excite mode 2. Thus the CMF or STR techniques are likely to fail when using the second mode if the source is at a depth of 26.5 m, and when using the third mode if the source is at a depth of 17 or 36 m.

This latter possibility is illustrated in Fig. 5, which shows the ranging results when a source is located 10 km from a 2° inclined array at a depth of 36 m (where mode 3 is not excited), with a received single-hydrophone SNR of 11.5 dB. The mode 1 and 2 pair generates a cross-correlation coefficient peak just above $\Delta_{12} \approx 7000$, but the mode 2 and 3 pair fails to generate a distinct CCC peak. Figure 5(b) shows the MSE from Eq. (9) for both mode pairs, (1,2) and (2,3). Because mode 3 was not excited, the (2,3)-mode pair result does not correctly range the source. However, both modes 1 and 2 have been excited by the source at this depth, so both ranging techniques are able to find the correct range using this mode combination.

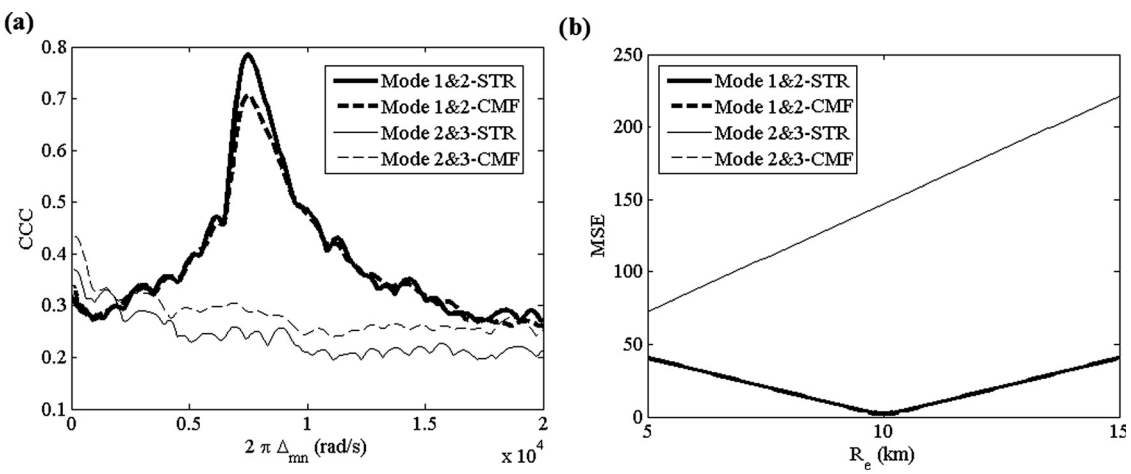

FIG. 5. Same as Fig. 4 except the simulated source signal is generated at 36 m depth. The (1,2) mode pair retains a minimum error at 10 km. However, because mode 3 was not excited by the sound source, the (2,3) mode pair is unable to estimate the correct range.

For the simulations shown in Figs. 4 and 5, both CMF and STR generate the same ranging results. By contrast, in Fig. 6 the sound source has been moved to 22 m depth (close to null of mode 2, 26.5 m) and has a smaller bandwidth (150–300 Hz). Because mode 2 has been only slightly excited in this simulation, both CMF and STR techniques have errors in ranging results (the estimated range by CMF and STR are 8.4 and 9.4 km, respectively). However, STR is more robust and estimates a more accurate range than CMF. We are thus forewarned that when processing real data, certain calling depths of the whale will yield values of Δ_{12} near zero, particularly at the zero of mode 2 (26.5 m). For whale calls close to this depth, inaccuracies in the range estimation will arise.

IV. ANALYSIS OF 2010 ARCTIC DEPLOYMENT

A. Equipment description

1. Vertical array

The data analyzed here were collected by an autonomous 16-hydrophone (or "element") vertical array built using two eight-channel array subsystems. Each subsystem used a cylindrical pressure case containing an eight-channel Persistor data acquisition system (CF2, Persistor Instruments, Inc., Marstons Mills MA), 4 Gb flash memory, and a 120 Gb hard drive. The pressure case also contained batteries and "HOBO" vertical inclinometers (Onset, Inc., Bourne, MA). From the top of the case, a custom Y-connector (Teledyne Impulse, San Diego, CA) branched into eight individual cables, each terminating in a HTI-94-SSQ hydrophone (High Tech Inc., Gulfport, MS) with −160 dB re 1 V/μPa sensitivity, and with 28 dB gain before A/D conversion. Acoustic data were sampled at an aggregate rate of 50 kHz, or 6.25 kHz per channel, with a gentle 35 dB roll-off between 1 and 2.5 kHz. During the time periods analyzed here, four of the channels had severe electrical noise, yielding 12 useable data channels. The array recorded for 2 h continuously to flash memory took 45 min to transfer data to hard disk, then waited an additional 2 h before starting the 43% duty cycle again.

The length of each branching cable was cut to yield 3 m hydrophone separations. All hydrophones, cables, and pressure cases were attached to 40 m of 1.25 cm diameter "superstrong" rope. One end of this rope was attached to a 32 kg Bruce anchor, which in turn was attached to 200 m of 3/8-in. crabbing line, terminated by a 4 kg Danforth anchor. The other thimbled end of the superstrong rope was attached to a cylindrically tapered foam subsurface float (50 cm long by 25 cm diameter) with 12 kg of buoyancy. A depth sensor was attached 2.6 m beneath the float.

2. DASARs

The DASARs record three acoustic signals continuously at a sampling rate of 1 kHz, using a suspended sensor protected by a spandex "sock" covering an aluminum frame. One channel samples data from an omnidirectional (pressure) hydrophone with a sensitivity of −149 dB re 1 V/μPa @ 100 Hz, with a high-pass frequency response of around 20 dB per decade to pre-whiten the expected ambient noise spectrum and thereby reduce the possibility of signal clipping arising from relatively high-amplitude low-frequency ambient noise. The other two channels measure acoustic particle velocity in orthogonal directions, with a sensitivity of 97 dB re 1 V/(m/s) at 100 Hz (which translates into −146 dB re 1 V/μPa at 100 Hz sensitivity for an acoustic plane wave). After a net 0 dB amplification, the data are converted to binary 16-bit samples, stored to flash memory, and then periodically dumped to a laptop hard drive. The electronic noise floor of the omnidirectional sensor is equivalent to a background noise level of 43 dB re 1 μPa^2/Hz at 100 Hz with a 6 dB/octave spectral slope. These values generally lay 15 dB below the Knudsen noise spectrum estimated for sea state 0. Specific details on the internal electronics are given in Greene *et al.* (2004), which describes an earlier version of a DASAR that uses a different sensor.

To determine the value of ϕ_{cal} for each DASAR [Eq. (12)], a series of calibrated transmissions were made

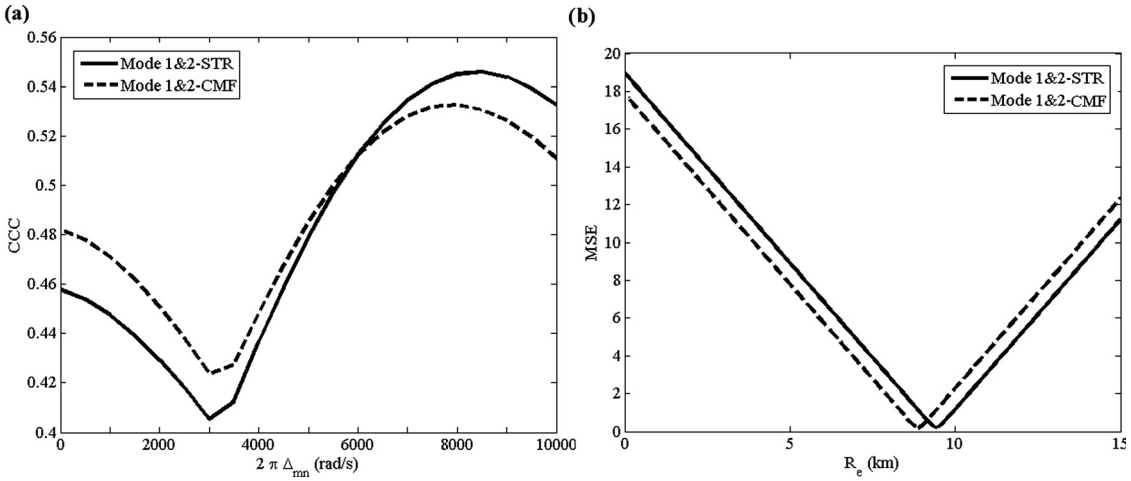

FIG. 6. (a) Cross-correlation coefficients vs Δ_{12} using STR (solid line) and CMF (dashed line). (b) Error calculated from Eq. (9) vs range R_e with CMF and STR. The broadcast signal is a 150–300 Hz chirp at 22 m depth and 10 km source-to-array range. The estimated ranges by CMF and STR are 8.4 and 9.4 km, respectively.

from three positions around each DASAR location. The projected waveform consisted of a 2-s tone at 400 Hz, a 2-s linear sweep from 400 to 200 Hz, a 2-s linear sweep from 200 to 400 Hz, a 2-s linear sweep from 400 to 200 Hz, and finally a 4-s long section of pseudo-random noise, i.e., an m-sequence with 255 chips, repeated once every second on a 255 Hz carrier frequency. A calibration is conducted just after a deployment and just before a pickup to confirm that the instrument did not rotate during the deployment.

B. Deployment geometry and experimental duration

The vertical array was deployed on 11 August 2010 at 70° 26.848′ N, 143° 12.572′ W, in 55 m water depth, and stopped recording on 6 September due to discharged batteries. The shallowest element was at 14.5 m (Fig. 1) and the deepest element was at 48.9 m, 6.1 m above the ocean floor. The tidal variation in the Beaufort Sea is generally small with less than 1 m variation. The element spacing was generally 3 m except for a portion of the array where the two subsystem cables overlapped, yielding six phones with 1.5 m spacing. Figure 1 illustrates the resulting overall array geometry. The measured array inclination was coarsely measured by inclinometers, which indicated that the array inclination was generally less than 5°, but the resolution of the sensors was too poor to provide any greater precision.

The array was deployed near Kaktovik, AK, in the vicinity of seven DASAR recorders (Greene *et al.*, 2004) arranged in a triangular grid (site 5 in Blackwell *et al.*, 2013). The array was roughly 4 km from the northernmost DASAR (G, see Fig. 7). The deployment depths at each DASAR indicate a gently sloping bottom ranging from 39 m at the southernmost DASAR (A) to 54 m at the deepest DASAR (G), and data from the Smith and Sandwell bathymetry database (Smith and Sandwell, 1997) confirm that the bottom bathymetry profile is relatively flat in directions parallel to the coast (Fig. 7). DASAR F was not recovered; hence six DASARs were available for triangulation. Fortunately, DASAR G was recovered, providing all of the spatial aperture that a full seven-DASAR array would provide. Thus the presence of a seventh DASAR would not be expected to dramatically improve the range resolution of the triangulation results reported here.

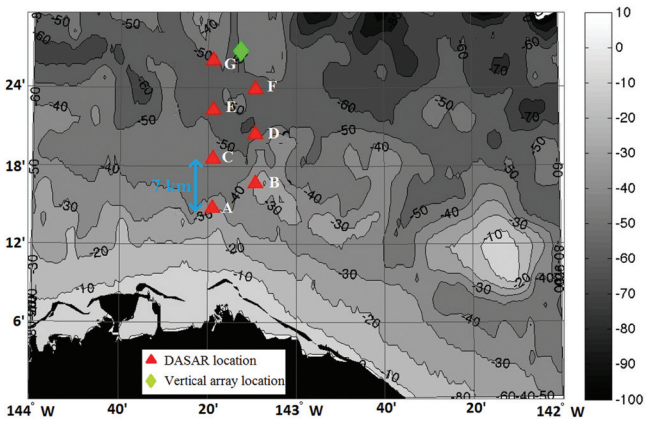

FIG. 7. Map of 2010 DASAR deployments in the Beaufort Sea. Triangles show DASAR locations, and the diamond shows the vertical array location. Bottom bathymetry (in meters) is obtained from Sandwell database. DASAR F was not operational during the 2010 array deployment.

C. Estimating the propagation environment and CMF/STR range uncertainty

Bowhead whale calls were first identified and localized from the DASAR recordings using an automated procedure (Thode *et al.*, 2012). To estimate the baseline propagation environment shown in Fig. 1, standard matched-field processing and geoacoustic inversion methods (Gerstoft, 1994; Gerstoft and Gingras, 1996; Thode *et al.*, 2000) were used on four broadband calls detected on August 20 and 31. Specifically, the Fourier coefficients from FFT snapshots of each array element were combined into a cross-spectral density matrix (CSDM) for every frequency component. Overlapping CSDM snapshots were then averaged to form a complete CSDM matrix from a transient call. For each frequency component, the ratio of the primary CSDM eigenvalue to the sum of the remaining eigenvalues was then used to estimate the SNR of the call at that frequency. Frequency components that yielded SNRs exceeding 10 dB were selected for combined MFP/geoacoustic inversions.

The geoacoustic inversions assumed a Pekeris waveguide model and solved for source range and depth; bottom speed, density, and attenuation; and array inclination. An optimization function was calculated by incoherently frequency-averaging Bartlett surfaces that were generated from the eigenvectors of the high-SNR call components and the output of a normal mode model. The optimization of this frequency-averaged objective function was achieved using the genetic inversion program SAGA (Gerstoft, 2007), combined with the range-independent SNAP normal mode model. SAGA created 40 individual inversion runs (or "populations") with each population consisting of 64 "individuals" with respective crossover, update, and mutation rates of 0.8, 0.5, and 0.05. Each inversion used 4000 iterations per population. All these optimization parameters were selected based on recommendations by Gerstoft (2007) and previous experience, and perturbing these parameters did not affect the results. The environmental parameter values from the inversion run that generated the highest value of the averaged Bartlett surface were selected as the final model for that call.

Table I lists the inverted range and bottom speed for the four calls chosen for the full-field geoacoustic inversions. The range of the calls varies from 1.2 to 37 km. The Pekeris model inverted from the 7.0 km range call in Table I was selected as the baseline environment for all subsequent simulations and data analyses because the median range of most automatically detected calls from the vertical array (using DASAR triangulation) was 7–8 km, and because the mean value of the bottom speed across all four inversions was close to the 7.0 km range value. The uncertainties in both range and bottom speed were estimated by taking the standard deviation of the 40 genetic algorithm runs for each call.

Figure 8 shows a specific example of a range-depth normalized Bartlett ambiguity surface, generated from a close-range broadband whale call measured on August 20 at

TABLE I. Summary of environmental and range parameters derived from the 40 MFP genetic algorithm Pekeris waveguide inversions described in Sec. IV C.

Call date and time	Best-fit range (gradient) ± standard deviation (km)	Bottom sound speed ± standard deviation uncertainty (m/s)	Normalized average Bartlett ambiguity value at optimized values. [Number of frequency components used]
20 Aug 2010, 01:42:39	1.2 ± 0.002	1707 ± 44	0.78 [5]
31 Aug 2010, 00:48:30	7.0 ± 0.02	1638 ± 42	0.82 [12]
31 Aug 2010, 10:50:13	17.7 ± 0.03	1691 ± 6	0.9 [12]
31 Aug 2010, 10:57:48	34.8 ± 0.6	1552 ± 37	0.8 [21]

01:42:39 (first row of Table I). The inversion used five frequency components between 45 and 253 Hz, and Fig. 8 uses the best-fit Pekeris environment described in Table I. The normalized incoherent frequency-averaged value at the mainlobe is 0.78 at a depth of 45 m and 1.2 km range, indicating that the environmental model is a good fit for this particular call.

Table I also offers insight into the uncertainty of the environmental parameters, which is the dominant source of error in modal interference range estimates like CMF and STR. The standard deviations of individual inversion results are small, but one sees that the bottom speed estimates across all four inversions have a spread of 5% around a mean value of 1647 m/s. The apparent bottom sound speed decreases with increasing source range; this might be an accurate representation of a mild range-dependent change in the ocean bottom, but may also indicate an increasing mismatch between the range-independent assumptions of the propagation model and an actual range-dependent bathymetry that shallows with increasing distance from the vertical array.

In subsequent sections, the uncertainty in CMF and STR range estimates will be defined as the cycle distance between the two modes used for ranging, evaluated at the center frequency of the bandwidth: $\Delta r = [2\pi/k_{mn}(\omega_c)]$. This number provides the distance to adjacent side lobes in the CMF and STR calculations. We have found that uncertainties in the environmental parameters can sometimes cause side lobes to yield the maximum CCC value, and thus the cycle distance is a conservative approximation of the range uncertainty arising from environmental uncertainty. At ranges between 5 and 10 km, this effective range uncertainty is between 15% and 20%. At higher ranges, this percentage error decreases.

D. Results of range comparisons

The experimental data displayed here were collected on August 31, 2010 between 00:46:15 and 10:57:48, a time period of low ambient noise, numerous whale calls, and a slightly earlier time period than what was used to estimate the propagation environment. The bowhead whales' east-west migration corridor is concentrated south of the array, in the vicinity of DASARs B and C in Fig. 7; hence no whale calls were detected closer than 7 km during this time.

A total of 19 bowhead whale calls were analyzed with frequencies in the range 50–500 Hz, which—at the low frequency end—is close to the cutoff frequencies of all propagating modes of the ocean sound channel. For all calls, the array was modeled as being perfectly vertical, with the exception of the last call (Call 19), which was found to yield

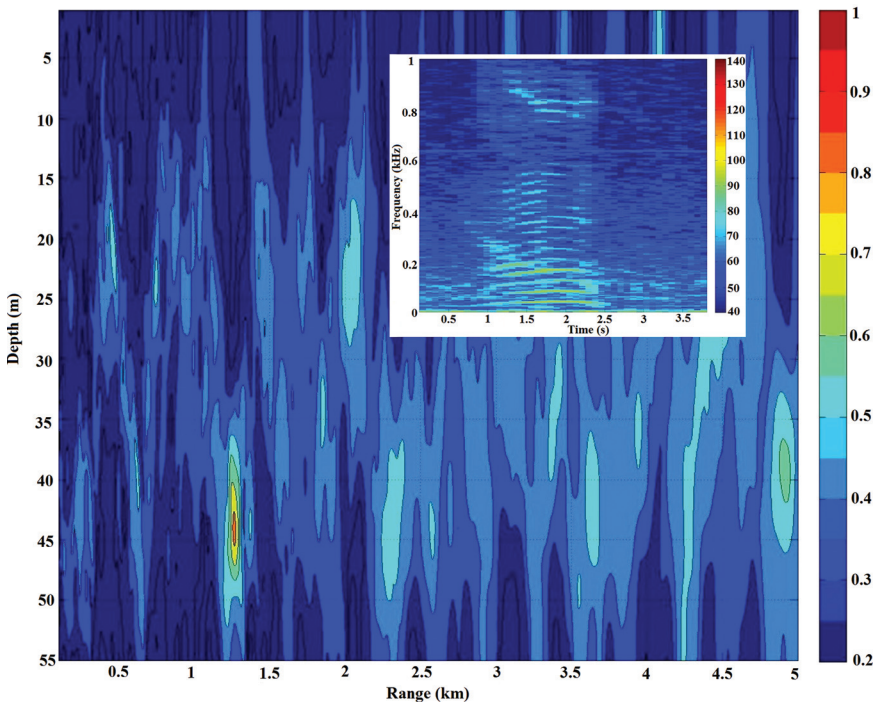

FIG. 8. Range-depth normalized Bartlett ambiguity surface for a broadband whale call measured on August 20 at 01:42:39 (inset), using the best-fit inverted Pekeris environment described in Table I. Six frequency components between 45 and 253 Hz have been incoherently averaged. The normalized mainlobe peak is 0.78 at a depth and range of 45 m and 1.2 km, respectively.

a range of zero when no inclination was modeled. A brute-force grid search of array inclinations between −5° and 5° yielded a non-zero range of 34.8 km for modeled inclinations between 2° and 3°. A matched-field inversion of this call (bottom row in Table I) for array inclination yielded 2.25°. As mentioned in Sec. IV B, the measured array inclination was less than 5° throughout most of the deployment. As discussed in the previous subsection, the inversion of the 7.5 km range call (second row, Table I) was selected for the STR and CMF processing, for two reasons: Most processed calls were near this range, and the inverted bottom speed for the 7.5 km call (1638 m/s) was close to the average of all inverted bottom speeds (1647 m/s).

Table II lists each call's bandwidth, SNR, range (i.e., distance between the estimated whale location and the vertical array) and range uncertainty using CMF, STR, and DASAR triangulation (plus estimated uncertainties), cross-correlation maximum values for the CMF and STR calculations, and the mode pair used for ranging. CMF and STR ranging techniques first attempted to use the first and second modes. If these modes were unsuccessful in ranging [i.e., maximizing Eq. (8) yields $\Delta_{12} = 0$], then the second and third mode pair was analyzed.

The DASAR ranges in Table II reflect the separation between the vertical array location and the maximum likelihood position derived from the DASAR triangulation. The orientation and size of the 90% uncertainty ellipse is unrelated to the relative position of the vertical array; instead, it arises from the relative position of the whale with respect to the DASAR distribution. To generate an estimate of range uncertainty for the DASAR results that can be directly compared to the CMF/STR uncertainties, the intersection of the 90% uncertainty ellipse with the line connecting the vertical array location and source was used. The procedure for estimating the uncertainties in the CMF and STR results has been discussed in the previous subsection; note that the resulting range uncertainties for both techniques in Table II are consistent with the full MFP inversions conducted in Table I, although Call 3 in Table II has a discrepancy of 500 m between the full-field Pekeris inversion and the STR/CMF result.

Figure 9 shows an example of the CMF and STR ranging results for a whale call recorded at 00:51:59 (Call 6 in Table II). In this figure, part (a) shows the spectrogram of the whale call between the two vertical lines; part (b) shows the cross-correlation coefficient between $\hat{S}_1(\omega)$ and $\hat{S}_2(\omega)e^{i2\pi\Delta_{12}/\omega}$ vs the dispersive-phase fitting constant Δ_{12}; and part (c) shows the MSE from Eq. (9) vs the estimated range R_e. The CMF and STR MSE curves are almost identical for this call. The ranges estimated by CMF and STR are 6.1 ± 1.3 km, and the DASAR-determined range is 7.2 ± 0.7 km.

Figure 10 shows the CMF and STR ranging results for a whale call recorded at 01:01:12 on August 31, 2010 (Call 14 in Table II) using the same format as Fig. 9. For this call, the (1,2)-mode pair was not successful in ranging (the best value of Δ_{12} was zero), so the (2,3)-mode pair was used, producing CMF- and STR-determined ranges of 10.1 ± 0.7 and 9.8 ± 0.7 km, respectively. The DASAR-triangulated range is 10.4 ± 2.4 km. For this call, the magnitude of the peak

TABLE II. Comparison between the performance of the CMF, STR, and DASAR techniques using experimental data from the Arctic. This table lists each call's bandwidth, SNR, ranges from CMF, STR, and DASAR (including estimated uncertainties), cross-correlation peak values for the CMF and STR calculations, and the mode pair used for successful ranging. The calls having underlined listings produced the same CMF- and STR-based ranging results. All call ranges were estimated using the propagation environment shown in Fig. 1. The vertical array inclination was assumed to be 0°, except for Call 19, which required a 2.25° inclination.

	Recorded time at the vertical array on 31 Aug 2010	Bandwidth (Hz)	SNR (dB)	CMF (km)	STR (km)	DASAR (km)	Peak CCC CMF (%)	Peak CCC STR (%)	Modes
1	00:46:15	160–270	20.7	8.1 ± 0.8	7.2 ± 0.8	7.3 ± 1.4	61.8	82.3	2&3
2	00:46:19	140–240	26.4	5.7 ± 0.7	6.7 ± 0.7	7.0 ± 1.6	59.7	68.4	2&3
3	00:48:30	50–320	17.2	8.7 ± 1.2	8.7 ± 1.2	7.2 ± 1.8	73.1	75.8	1&2
4	00:49:04	140–280	17.8	9.5 ± 1.3	8.3 ± 1.3	7.2 ± 2.4	90.3	91.3	1&2
5	00:49:19	170–290	19.2	9.1 ± 1.4	7.9 ± 1.4	7.0 ± 2.4	78.5	85.7	1&2
6	00:51:59	100–320	23.8	6.1 ± 1.3	6.1 ± 1.3	7.2 ± 0.7	85.1	91.5	1&2
7	00:53:31	160–320	20.3	0.0 ± 1.5	6.8 ± 1.5	7.3 ± 2.0	53.8	58.9	1&2
8	00:54:36	80–180	20.1	6.7 ± 0.7	6.7 ± 0.7	7.5 ± 1.8	64.3	67.5	1&2
9	00:54:48	140–280	17.4	8.3 ± 1.3	7.6 ± 1.3	7.4 ± 2.0	55.3	62.3	1&2
10	00:54:52	100–180	18.7	8.6 ± 0.9	8.6 ± 0.9	7.5 ± 1.5	88.2	94.6	1&2
11	00:54:54	70–300	16.4	8.2 ± 1.2	8.2 ± 1.2	7.5 ± 1.6	79.1	84.8	1&2
12	00:58:11	40–100	7.0	8.0 ± 0.5	8.0 ± 0.5	7.6 ± 1.0	78.3	88.2	1&2
13	01:00:56	80–420	8.0	8.5 ± 0.9	8.5 ± 0.9	7.1 ± 1.0	37.7	44.1	2&3
14	01:01:12	140–240	12.4	10.1 ± 0.7	9.8 ± 0.7	10.4 ± 2.4	68.5	78.7	2&3
15	01:24:19	180–400	16.1	18.2 ± 1.8	16.5 ± 1.8	16.0 ± 2.8	46.7	60.5	1&2
16	01:26:27	140–340	8.2	15.5 ± 1.5	14.3 ± 1.5	15.7 ± 0.2	62.8	77.3	1&2
17	01:28:26	170–250	16.7	16.5 ± 0.8	16.5 ± 0.8	15.7 ± 0.2	75.4	92.6	2&3
18	01:38:31	120–280	11.5	28.7 ± 1.3	28.7 ± 1.3	24.5 ± 1.2	78.5	84.1	1&2
19	10:57:48	100–170	10.9	36.6 ± 0.9	37.9 ± 0.9	40.7 ± 10.7	63.0	65.1	1&2

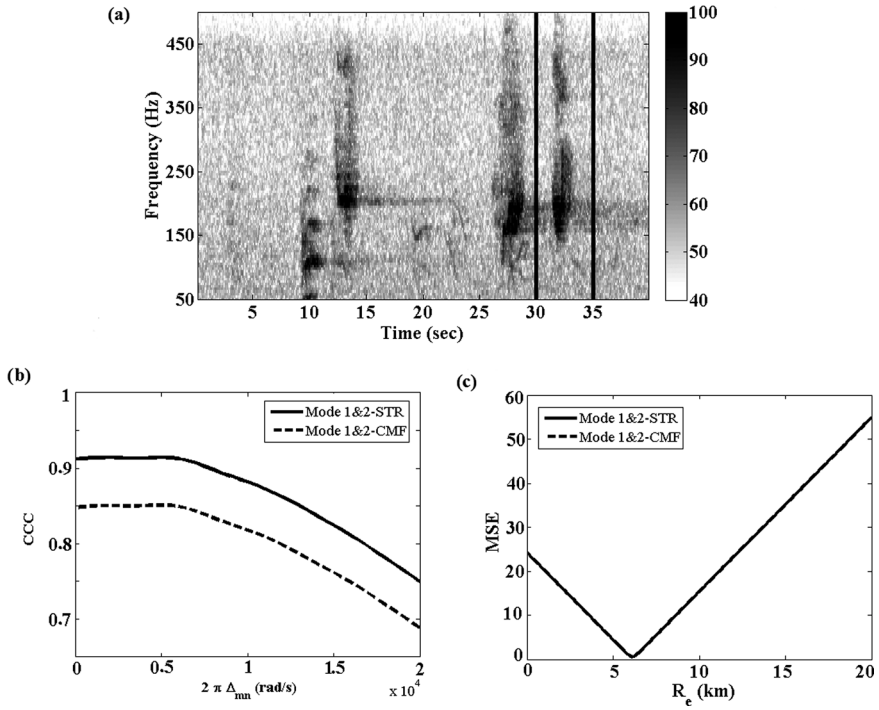

FIG. 9. (a) Spectrogram of bowhead call at 00:51:59 (row 6, Table II); (b) cross-correlation coefficient vs Δ_{12} using (1,2) mode pair for CMF (dashed line) and STR (solid line), (c) error calculated from Eq. (9) vs range using (1,2) mode pair with CMF and STR calculation. The CMF- and STR-determined ranges for this call are both 6.1 ± 1.3 km, while the DASAR-determined range is 7.2 ± 0.7 km.

cross-correlation coefficient presented in Fig. 10(b) is lower than that shown in Fig. 9(b); possible reasons for this are that a lower bandwidth is available, different mode pairs have been used, or the propagation environment is gradually changing at 10 km, increasing the environmental mismatch between the modeled and measured dispersion properties.

Figure 11 displays a comparison between the STR and DASAR range estimates for four different whale calls tabulated in Table II, along with the uncertainties associated with both. Figure 11(d) also illustrates the definitions of the DASAR range and DASAR range uncertainty. One sees that the STR ranges are consistent with the DASAR estimates to within the uncertainties of both measurements. Also note that the DASAR triangulation results degrade rapidly once the source lies outside the array geometry [Fig. 11(d)].

E. Discussion of experimental results

The nine calls having underlined listings in Table II produced the same CMF- and STR-based ranging results. Overall, the vertical-array range estimates are consistent with the DASAR triangulation results and apparently are robust out to ranges of at least 35 km. Viable whale call range estimates were produced by both the CMF- and STR-based approaches even when the CCC values fall below 40%. Note that potential biases from range-dependent

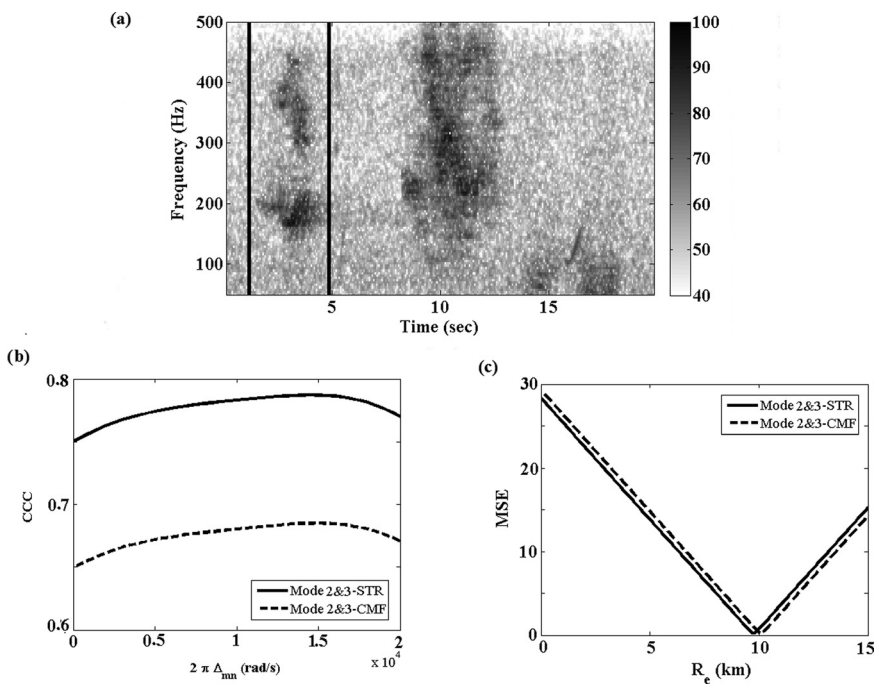

FIG. 10. (a) Spectrogram at 01:01:12 (row 14, Table II); (b) cross-correlation coefficient vs Δ_{23} using (2,3) mode pair for CMF (dashed line) and STR (solid line); (c) error calculated from Eq. (9) vs range using (2,3) mode pair with CMF and STR calculation. The DASAR-determined range is 10.4 ± 2.4 km for this whale call. CMF- and STR-determined ranges are 10.1 ± 0.7 and 9.8 ± 0.7 km, respectively.

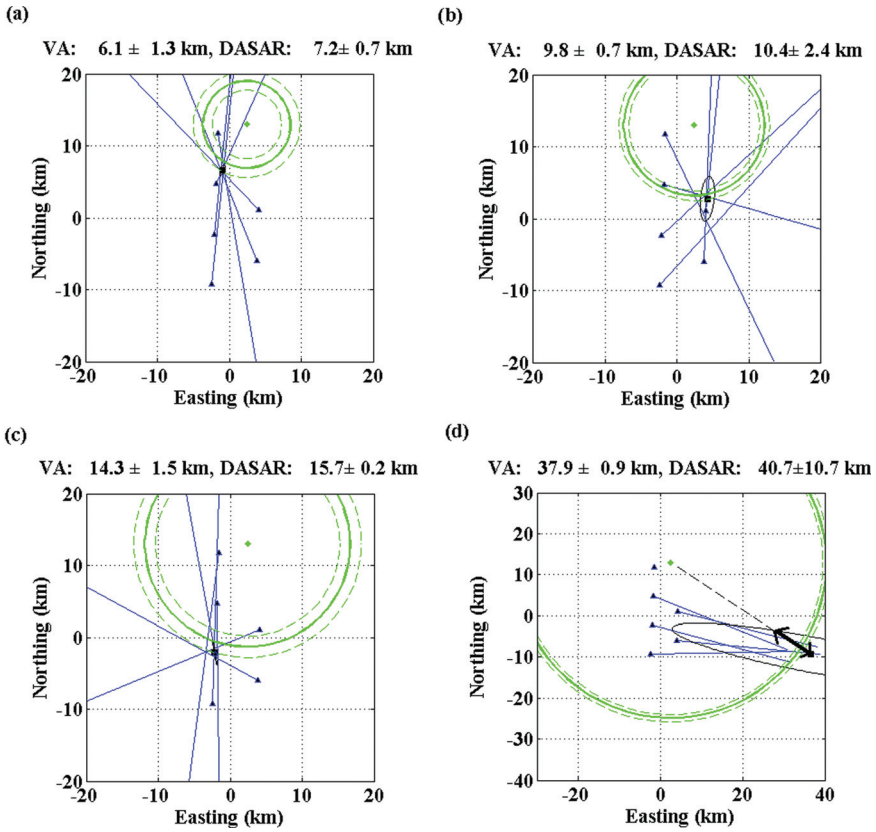

FIG. 11. Comparison between STR and DASAR ranging techniques for four sample whale calls (a) 00:51:59 (row 6, Table II); (b) 01:01:12 (row 14); (c) 01:26:27 (row 16); (d) 10:57:48 (row 19). DASAR and vertical array locations are shown by triangles and a diamond, respectively. Solid thin dark lines show the DASAR bearing estimates and 90% confidence ellipse, while the solid circle shows STR ranging results with range uncertainties estimated by the two-mode cycle distances (dashed circles). The thin dashed line in (d) illustrates the definition of "DASAR range," while the double arrow indicates the DASAR range uncertainty listed in Table II.

environment changes are not accounted for here; however, if such biases exist, they appear to lie within the range uncertainties of the triangulation estimates. The fact that the dispersion-based range estimates are accurate past 30 km range seems surprising, given the fact that the bathymetry map in Fig. 7 suggests that the water depth shallows from 55 m to under 30 m to the southeast of the array. However, the precision of the DASAR range estimates deteriorates even more quickly at greater ranges, as errors and biases in the bearing estimates become magnified for locations far outside the distributed DASAR array.

Figure 12 shows some details for the most distant example in Table II (Call 19). If the array was modeled without any inclination (thin solid and dashed lines), then the predicted range was zero because Eq. (3) is invalid, and the mode 2 signal estimate becomes heavily contaminated by the mode 1 contribution, yielding a maximum cross-correlation at zero time lag in Eq. (6). When a 2.25° inclination is included in the propagation model (thick lines), both CMF and STR techniques yield a non-zero result that is consistent with the maximum-likelihood triangulation estimate. This example shows that the vertical array-based ranging techniques can be sensitive to mismatches in array inclination; however, the result of a mismatched inclination is typically range estimates of zero for all mode pairs, so flaws in the inclination estimate can be recognized and corrected.

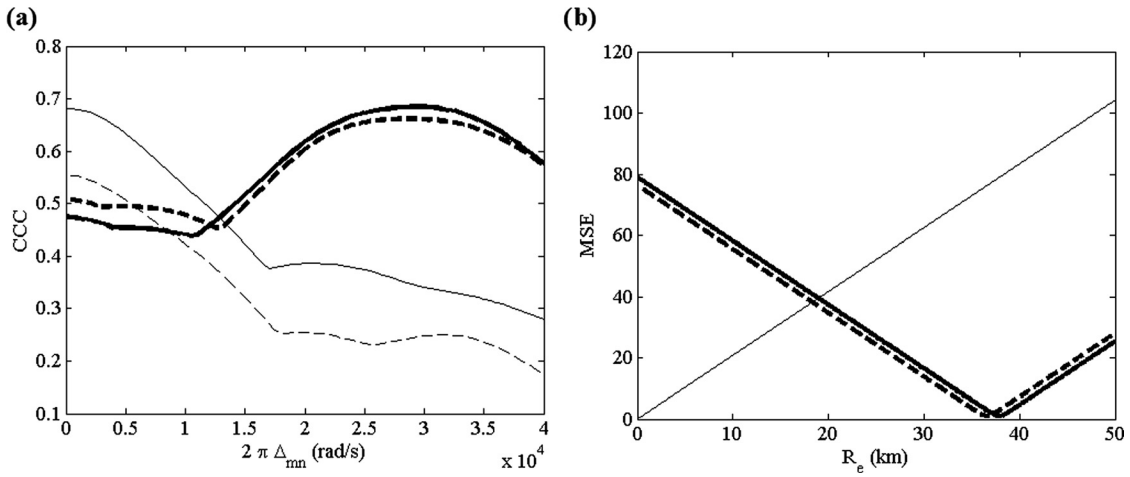

FIG. 12. (a) Cross-correlation coefficient vs Δ_{12} for Call 19 in Table II. Thin and thick lines show respective results for a 0° and 2.25° modeled array inclination, while STR and CMF results are shown by solid and dashed lines, respectively. (b) Error calculated from Eq. (9) vs estimated range using (1,2) mode pair.

Based on the results in Table II, the STR-based vertical-array ranging technique appears to be slightly superior to the CMF-based technique for at least two reasons. First, it consistently produces a higher peak CCC value and did not fail to adequately range a single call. Ranging based on the mode filtering technique failed once (Call 7 in Table II). Figure 13 shows the cross-correlation coefficient vs Δ_{mn} plot for Call 7. The broad maxima and relatively small difference between the peak CCC value and other CCC values for both techniques indicate this is a marginal case for ranging and suggest that the whale that made this call may have been near the node of the second mode. However, the STR technique seems to have succeeded in this difficult case.

When all 19 calls are considered together, there is also a slight difference in the accuracy of the two vertical-array ranging techniques, if one assumes that the triangulation ranges are accurate. The root-mean square (rms) distance difference, R_{rms}, between the estimated ranges, R_e, and DASAR-determined ranges, R_{DASAR}, were computed from

$$R_{rms} = \left[\frac{1}{19}\sum_{n=1}^{19}(R_e - R_{DASAR})_n^2\right]^{1/2} \quad (14)$$

for both vertical array techniques. The rms differences from the DASAR triangulations are 0.56 and 0.32 km for the CMF-based and STR-based techniques, respectively. These values are less than 5% and 3% of the average (DASAR) range for the 19 calls, 11.4 km. Thus the STR ranging technique seems slightly better than CMF technique when compared with DASAR ranging results in these low ambient-noise conditions. When Call 7 is removed, the CMF-determined rms range error for the remaining 18 calls improves to 0.43 km. In either case, these ranging differences are most likely the result of minor uncorrected array tilt (Calls 1–18) and violations of the environmental range-independence assumption inherent in the CMF and STR ranging approaches.

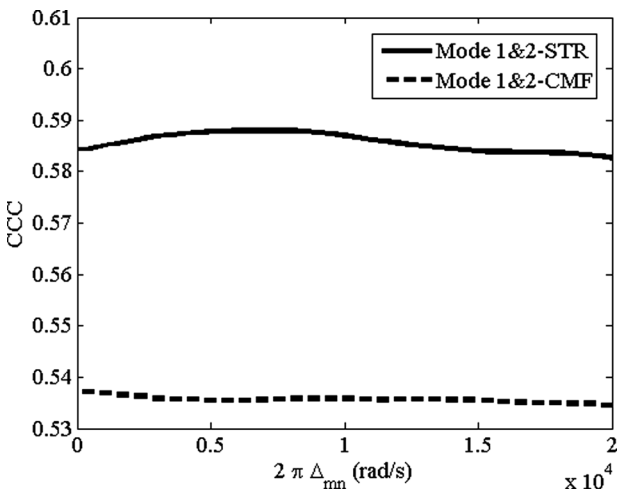

FIG. 13. Cross-correlation coefficient vs Δ_{12} using (1,2) mode pair for Call 7 in Table II. CMF and STR results are shown by the dashed line and the solid line, respectively. This is a marginal case for ranging a whale call successfully. STR has slightly higher peak cross-correlation coefficient than CMF, and it is able to estimate the source range while CMF does not.

V. CONCLUSION

The source ranging performance of two dispersion-based techniques, CMF and mode-based STR, has been investigated and compared to triangulation results using simulations and experimental data from the Arctic Ocean under conditions where modal dispersion was an important characteristic of the signals. The simulations mimicked the signals and geometry of the Arctic experiment, and the ranging performance of both techniques was found to be good, even nearly identical, in simulations when both modes are completely excited. However, the STR technique performs slightly better in simulation whenever one mode is only marginally excited by the sound source. The STR technique was also more robust and more accurate when applied to 19 high SNR bowhead whale calls culled from experimental data from the Arctic Ocean environment. For example, STR successfully produced estimated ranges for all calls while CMF failed for one call. The failure for that call most likely occurred because the depth where the call originated prevented adequate excitation of one of the modes used for ranging. STR results also compared well with the triangulation results obtained from using a full horizontally-distributed array and even displayed more precision for calls originating well outside the distributed array boundaries. Neither range-dependent bathymetric effects nor mismatches in vertical array inclination proved to be insurmountable obstacles for CMF or STR. However, the performance of all three systems under lower SNR conditions remains to be examined.

These findings may be important for passive remote study of marine mammals in that they may reduce the cost and field effort of monitoring large regions adjacent to potentially disturbing anthropogenic noise sources. Application scenarios for this technique involve Navy active sonar testing, seismic surveying, and pile-driving in lakes, rivers, and oceans. For successful ranging using the techniques developed here, whale calls should have at least 50 Hz bandwidth and occupy a bandwidth where the phase velocities of at least two propagating modes are different. If calls occupy a frequency band too far from the cut-off frequencies, the phase velocities of all modes at all frequencies will be very similar, so the reconstructed signals from mode 1 and 2 will be the same and the constant Δ_{mn} in Eq. (7) cannot be reliably found. Figure 2 suggests that this situation will not happen until at least 300 Hz for the Arctic environments shown here; this suggests that some pinniped calls may be ranged with these techniques as well.

ACKNOWLEDGMENTS

This research effort was supported by the Office of Naval Research under Grant No. N00014-11-1-0047. The vertical array data collection was supported by the North Pacific Research Board, Project 496. Shell Exploration and Production Company (SEPCO) facilitated the deployment and retrieval of the arrays and provided access to the data from the DASAR array, Greeneridge Sciences Inc. (GS) allowed the use of the DASAR data for whale call range confirmation, and Fairweather Inc. helped arrange the logistics

of the deployments. The authors also thank Delphine Mathias and the entire crew of the R/V Norseman II for their outstanding help in the field, with particular thanks to deck chief "Scotty" Hameister, Charles R. Greene, Jr. of GS and A. Michael Macrander of SEPCO for providing steady support for the array deployments over three years, including 2010, the year addressed in this study.

Abadi, S. H., Rouseff, D., and Dowling, D. R. (**2012**). "Blind deconvolution for robust signal estimation and approximate source localization," J. Acoust. Soc. Am. **131**(4), 2599–2610.

Baggeroer, A. B., Kuperman, W. A., and Mikhalevsky, P. N. (**1993**). "An overview of matched- field processing methods in ocean acoustics," IEEE J. Ocean. Eng. **18**, 401–424.

Blackwell, S. B., Nations, C. S., McDonald, T. L., Greene, C. R., Thode, A. M., Guerra, M., and Macrander, A. M. (**2013**). "Effects of airgun sounds on bowhead whale calling rates in the Alaskan Beaufort Sea," Marine Mam. Sci. **29**, E342–E365.

Buck, J. R., Preisig, J. C., and Wage, K. E. (**1998**). "A unified framework for mode filtering and the maximum a posteriori mode filter," J. Acoust. Soc. Am. **103**, 1813–1824.

Bucker, H. P. (**1976**). "Use of calculated sound fields and matched-field detection to locate sound sources in shallow water," J. Acoust. Soc. Am. **59**, 368–373.

Clark, C. W., and Ellison, W. T. (**2000**). "Calibration and comparison of the acoustic location methods used during the spring migration of the bowhead whale, *Balaena mysticetus*, off Pt. Barrow, Alaska, 1984–1993," J. Acoust. Soc. Am. **107**(6), 3509–3517.

Clark, C. W., and Johnson, J. H. (**1984**). "The sounds of the bowhead whale, *Balaena mysticetus*, during the spring migrations of 1979 and 1980," Can. J. Zool. **62**, 1436–1441.

Collison, N. E., and Dosso, S. E. (**2000**). "Regularized matched-mode processing for source localization," J. Acoust. Soc. Am. **107**, 3089–3100.

Cummings, W. C., and Holliday, D. V. (**1987**). "Sounds and source levels from bowhead whales off Pt. Barrow, Alaska," J. Acoust. Soc. Am. **82**, 814–821.

D'Spain, G. L., Kuperman, W. A., Clark, C. W., and Mellinger, D. K. (**1995**). "Simultaneous source ranging and bottom geoacoustic inversion using shallow water, broadband dispersion of fin whale calls," J. Acoust. Soc. Am. **97**, 3353.

Gerstoft, P. (**1994**). "Global inversion by genetic algorithms for both source position and environmental parameters," J. Comp. Acoust. **2**, 251–266.

Gerstoft, P. (**2007**). "SAGA Users guide 5.4, an inversion software package," (2007).

Gerstoft, P., and Gingras, D. F. (**1996**). "Parameter estimation using multi-frequency range– dependent acoustic data in shallow water," J. Acoust. Soc. Am. **99**, 2839–2851.

Grachev, G. A. (**1993**). "Theory of acoustic field invariants in layered waveguides," Acoust. Phys. **39**, 67–71.

Greene, C. R., McLennan, M. W., Norman, R. G., McDonald, T. L., Jakubczak, R. S., and Richardson, W. J. (**2004**). "Directional frequency and recording (DIFAR) sensors in seafloor recorders to locate calling bow-head whales during their fall migration," J. Acoust. Soc. Am. **116**, 799–813.

Huber, P. J. (**1964**). "Robust estimation of a location parameter," Ann. Math. Stat. **35**, 73–78.

Jensen, F., Kuperman, W., Porter, M., and Schmidt, H. (**1994**). *Computational Ocean Acoustics* (AIP, New York).

Krolik, J. L. (**1992**). "Matched-field minimum variance beamforming in a random ocean channel," J. Acoust. Soc. Am. **92**(3), 1408–1419.

Lee, S., and Makris, N. C. (**2006**). "The array invariant," J. Acoust. Soc. Am. **119**, 336–351.

Lenth, R. V. (**1981a**). "On finding the source of a signal," Technometrics **23**, 149–154.

Lenth, R. V. (**1981b**). "Robust measures of location for directional data," Technometrics **23**, 77–81.

Lin, Y. T., Newhall, A. E., and Lynch, J. F. (**2012**). "Low-frequency broadband sound source localization using an adaptive normal mode back-propagation approach in a shallow-water ocean," J. Acoust. Soc. Am. **131**, 1798–1813.

Lo, E., Zhou, J., and Shang, E. (**1983**). "Normal mode filtering in shallow water," J. Acoust. Soc. Am. **74**(6), 1833–1836.

Lu, I. T., Chen, H. Y., and Voltz, P. (**1993**). "A matched-mode processing technique for localizing a transient source in the time domain," J. Acoust. Soc. Am. **93**, 1365–1373.

Mathias, D., Thode, A. M., Straley, J. and Andrews. R. D. (**2013**). "Acoustic tracking of sperm whales in the Gulf of Alaska using a two-element vertical array and tags," J. Acoust. Soc. Am. **134**, 2446–2461.

Moore, S. E., Stafford, K. M., Mellinger, D. K., and Hildebrand, J. A. (**2006**). "Listening for large whales in the offshore waters of Alaska," Bioscience **56**, 49–55.

Morrissey, R. P., Ward, J., DiMarzio, N., Jarvis, S., and Moretti, D. J. (**2006**). "Passive acoustic detection and localization of sperm whales (*Physeter macrocephalus*) in the tongue of the ocean," Appl. Acoust. **67**(11–12), 1091–1105.

Munger, L. M., Wiggins, S. M., and Hildebrand, J. A. (**2011**). "North Pacific right whale up-call source levels and propagation distance on the southeastern Bering Sea shelf," J. Acoust. Soc. Am. **129**, 4047–4054.

Nosal, E. M. (**2013**). "Methods for tracking multiple marine mammals with wide-baseline passive acoustic arrays," J. Acoust. Soc. Am. **134**, 2383–2392.

Nosal, E. M., and Frazer, L. N. (**2007**). "Modified pairwise spectrogram processing for localization of unknown broadband sources," IEEE J. Ocean. Eng. **32**, 721–728.

Porter, M. B., and Reiss, E. L. (**1984**). "A numerical method for ocean acoustic normal modes," J. Acoust. Soc. Am. **76**, 244–252.

Ratilal, P., Gerstoft, P., Pong, Y. K., and Goh, J. T. (**1998**). "Inversion of pressure data on a vertical array for seafloor geoacoustic properties," J. Comp. Acoust. **6**, 269–289.

Richardson, W. J., Würsig, B., and Greene, C. R. (**1986**). "Reactions of bowhead whales, *Balaena-mysticetus*, to seismic exploration in the Canadian Beaufort Sea," J. Acoust. Soc. Am. **79**, 1117–1128.

Rideout, B. P., Dosso, S. E., and Hannay, D. E. (**2013**). "Underwater passive acoustic localization of Pacific walruses in the northeastern Chukchi Sea," J. Acoust. Soc. Am. **134**, 2534–2545.

Sabra, K. G., and Dowling, D. R. (**2004**). "Blind deconvolution in oceanic waveguides using artificial time reversal," J. Acoust. Soc. Am. **116**, 262–271.

Sabra, K. G., Song, H. C., and Dowling, D. R. (**2010**). "Ray-based blind deconvolution in ocean sound channels," J. Acoust. Soc. Am. **127**, EL42–EL47.

Shang, E. C. (**1989**). "An efficient high-resolution method of source localization processing in mode space," J. Acoust. Soc. Am. **86**, 1960–1964.

Shang, E. C., Clay, C. S., and Wang, Y. Y. (**1985**). "Passive harmonic source ranging in waveguides by using mode filter," J. Acoust. Soc. Am. **78**, 172–175.

Smith, W. H., and Sandwell, D. T. (**1997**). "Global seafloor topography from satellite altimetry and ship depth soundings," Science **277**, 1956–1962.

Thode, A. M., D'Spain, G. L., and Kuperman, W. A. (**2000**). "Matched-field processing and geoacoustic inversion of blue whale vocalizations," J. Acoust. Soc. Am. **107**, 1286–1300.

Thode, A. M., Gerstoft, P., Burgess, W. C., Sabra, K. G., Guerra, M., Stokes, D. M., Noad, M., and Cato, D. H. (**2006**). "A portable matched-field processing system using passive acoustic time synchronization," IEEE J. Ocean. Eng. **31**, 696–710.

Thode, A. M., Kim, K. H., Blackwell, S. B., Greene, C. R., McDonald, T. L., and Macrander, M. A. (**2012**). "Automated detection and localization of bowhead whale sounds in the presence of seismic airgun surveys," J. Acoust. Soc. Am. **131**, 3726–3747.

Tiemann, C. O., Porter, M. B., and Frazer, L. N. (**2004**). "Localization of marine mammals near Hawaii using an acoustic propagation model," J. Acoust. Soc. Am. **115**, 2834–2843.

Tiemann, C. O., Thode, A. M., Straley, J., O'Connell, V., and Folkert, K. (**2006**). "Three-dimensional localization of sperm whales using a single hydrophone," J. Acoust. Soc. Am. **120**, 2355–2365.

Valtierra, R. D., Holt, R. G., Cholewiak, D., and Van Parijs, S. M. (**2013**). "Calling depths of baleen whales from single sensor data; development of an autocorrelation method using multipath localization," J. Acoust. Soc. Am. **134**, 2571–2581.

von Benda-Beckmann, A. M., Beerens, S. P., and van IJsselmuide, S. P. (**2013**). "Effect of towed array stability on instantaneous localization of marine mammals," J. Acoust. Soc. Am. **134**, 2409–2417.

Wiggins, S. M., Munger, L. M., Moore, S. E., and Hildebrand, J. A. (**2004**). "Waveguide propagation allows range estimates for North Pacific right whales in the Bering Sea," Can. Acoust. **32**, 146–154.

Wilcock, W. S. (**2012**). "Tracking fin whales in the northeast Pacific Ocean with a seafloor seismic network," J. Acoust. Soc. Am. **132**(4), 2408–2419.
Yang, T. C. (**1987**). "A method of range and depth estimation by modal decomposition," J. Acoust. Soc. Am. **82**, 1736–1745.
Yang, T. C. (**1989**). "Effectiveness of mode filtering: A comparison of matched-field and matched-mode processing," J. Acoust. Soc. Am. **87**(5), 2072–2084.
Yoo, K., and Yang, T. C. (**1999**). "Broadband source localization in shallow water in the presence of internal waves," J. Acoust. Soc. Am. **106**, 3255–3269.
Zimmer, W. M. X. (**2013**). "Range estimation of cetaceans with compact volumetric arrays," J. Acoust. Soc. Am. **134**, 2610–2618.
Zurk, L. M., and Ward, J. (**2003**). "Source motion mitigation for adaptive matched field processing," J. Acoust. Soc. Am. **113**, 2719–2731.