ICES Journal of Marine Science

ICES Journal of Marine Science (2018), doi:10.1093/icesjms/fsy127

Temporal–spatial, spectral, and source level distributions of fin whale vocalizations in the Norwegian Sea observed with a coherent hydrophone array

Heriberto A. Garcia[1], Chenyang Zhu[1], Matthew E. Schinault[1], Anna I. Kaplan[1], Nils Olav Handegard[2], Olav Rune Godø[2], Heidi Ahonen[3], Nicholas C. Makris[4], Delin Wang[1], Wei Huang[1], and Purnima Ratilal[1]*

[1]*Laboratory for Ocean Acoustics and Ecosystem Sensing, Northeastern University, 360 Huntington Avenue, Boston, Massachusetts 02115, USA*
[2]*Institute of Marine Research, Post Office Box 1870, Nordnes, N-5817 Bergen, Norway*
[3]*Norwegian Polar Institute, N-9296 Tromsø, Norway*
[4]*Laboratory for Undersea Remote Sensing, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA*

**Corresponding author: tel: 617-373-8458; e-mail: purnima@ece.neu.edu*

Garcia, H. A., Zhu, C., Schinault, M. E., Kaplan, A. I., Handegard, N. O., Godø, O. R., Ahonen, H., Makris, N. C., Wang, D., Huang, W., and Ratilal, P. Temporal–spatial, spectral, and source level distributions of fin whale vocalizations in the Norwegian Sea observed with a coherent hydrophone array. – ICES Journal of Marine Science, doi:10.1093/icesjms/fsy127.

Received 16 February 2018; revised 30 June 2018; accepted 26 July 2018.

To better understand fin whale vocalization behaviour in the Norwegian and Barents Seas, a large-aperture densely sampled coherent hydrophone array was deployed in late winter 2014 to monitor their vocalizations instantaneously over wide areas via passive ocean acoustic waveguide remote sensing (POAWRS). Here, we (i) provide a time-frequency characterization for different call types observed (20 Hz pulses, 130 Hz upsweeps, 30–100 Hz downsweep chirps, and 18–19 Hz backbeats); (ii) compare their relative abundances in three different coastal regions off Alesund, Lofoten, and Northern Finnmark; (iii) estimate the temporal and spatial distributions; (iv) source level distributions; and (v) probability of detection (PoD) regions for the more abundant 20 Hz pulse and 130 Hz upsweep call types. The fin whale vocalizations received over the diel cycle (24 h) were significantly more abundant by a factor of roughly seven off Northern Finnmark than the other two regions, associated with fish feeding activities. The source levels are estimated to be 190.5±7.4 dB for the fin whale 20 Hz pulses and 170.3 ± 5.2 dB for the 130 Hz upsweeps. We find that fin whales are capable of producing each vocalization type either independently or simultaneously with other types, and the 20 Hz sound production in the fin whales involves a mechanism that generates a significantly less-intense second-order harmonic of the fundamental.

Keywords: 20 Hz, 130 Hz, beamforming, detection range, directional sensing, fin whale, localization, passive ocean acoustic waveguide remote sensing, passive sensing, probability of detection, source level, vocalization

Introduction

The vocalization behaviour of fin whales (*Balaenoptera physalus*) in the Norwegian and Barents Seas is monitored and studied using a large-aperture, densely-sampled coherent hydrophone array system with 160 hydrophone elements. The passive ocean acoustic waveguide remote sensing (POAWRS) technique is employed to provide detection, bearing-time estimation, time-frequency characterization, and classification, as well as localization and geographic positioning of the fin whale vocalizations received instantaneously over wide areas greater than 10 000 km^2. The observations were made from 18 February to 8 March 2014 in several regions of the Norwegian and Barents Seas (Figure 1) coinciding with the spawning season and grounds for three commercially and ecologically important fish species; the Atlantic herring

(*Clupea harengus*) off the coast of Alesund, the Atlantic cod (*Gadus morhua*) off the Lofoten archipelago, and the capelin (*Mallotus villosus*) off the Northern Finnmark coast.

Atlantic herring, cod, and capelin all spawn along the northern Norwegian coast during spring time, which presents an enormous transport of biomass from oceanic to coastal locations (Nakken, 2008). The release of biomass as spawning product is of importance to coastal ecology. Capelin spawn at the northern Norwegian coast during February–March and concentrate in coastal areas well suited for predation. Cod feed partly on capelin during their migration to their more southern spawning areas concentrated around the Lofoten Islands, with main spawning activity occurring in March–April. Herring migrate from various overwintering locations in the north to spawning areas all along the western Norwegian coast, in recent years concentrated in the Møre area. For marine mammals that are top predators, such as the fin whale, the concentrated fish migrations and spawnings are a tremendous source of prey. It is unclear how the recovery of large baleen whales will impact large oceanic fish stocks in terms of future harvesting potential. It is therefore of crucial importance to develop methodologies to observe marine mammals over wide areas and gather the information required to understand their behaviour, including their interaction with fish species.

Due to the limitations of visual sightings above water, passive acoustic monitoring of fin whale vocalizations has become an important method for investigation of fin whale sound production and behaviour underwater. The vocalizations of the fin whale have been studied for several offshore environments of the world, including various coastal and ocean regions of the Atlantic (Watkins *et al.*, 1987; Edds, 1988; Clark and Gagnon, 2004; Nieukirk *et al.*, 2004; Simon *et al.*, 2010; Klinck *et al.*, 2012; Wang *et al.*, 2016a) and the Pacific (Northrop *et al.*, 1968; Thompson *et al.*, 1992; McDonald *et al.*, 1995; Charif *et al.*, 2002; Weirathmueller *et al.*, 2013), the Southern Ocean (Širović *et al.*, 2007), and the Mediterranean Sea (Clark *et al.*, 2002; Castellote *et al.*, 2012). The 20 Hz pulse vocalization has been found to be ubiquitous for fin whales in all ocean regions studied previously. Based on observations of a sample of fin whale individuals during mating, repetitive bouts of the 20 Hz pulses were found to be vocalized by male fin whales (Watkins *et al.*, 1987; Croll *et al.*, 2002). More randomly occurring fin whale 20 Hz pulse vocalizations have been associated with other communication purposes, such as serving as contact signals for coordinated activities during feeding and migration (McDonald *et al.*, 1995; Wang *et al.*, 2016a). The other fin whale vocalization types are not as common (Castellote *et al.*, 2012) since their observations are highly specific to certain ocean regions, with measurement rates dependent on the receiver type.

The time–frequency characteristics of fin whale 20 Hz pulse vocalizations have been quantified in previous studies, providing information on parameters such as peak frequency, duration, and inter-pulse intervals (IPIs) for repetitive bouts of this vocalization type (Castellote *et al.*, 2012). These characteristics of the 20 Hz pulse vocalizations, and the presence or absence of other fin whale call types have been found to be useful for inferring population structure (Castellote *et al.*, 2012). The 20 Hz pulse vocalization source level (Watkins *et al.*, 1987; Charif *et al.*, 2002; Širović *et al.*, 2007; Weirathmueller *et al.*, 2013; Wang *et al.*, 2016b) has been estimated for fin whales in several different offshore regions. In general, the fin whale 20 Hz pulse vocalizations have been found to be highly intense, accounting for their consistent

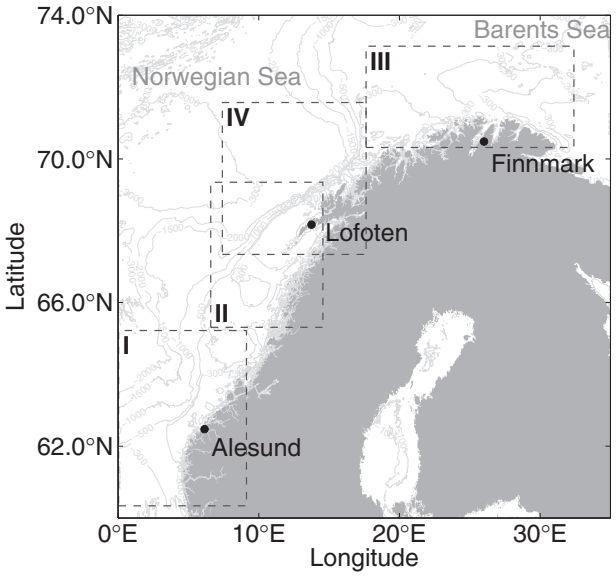

Figure 1. Observation regions of the coherent hydrophone array during the Norwegian Sea Experiment 2014 from 18 February to 8 March. The offshore regions off Alesund (I), Lofoten (II and IV), and the Northern Finnmark (III) are shown by the dotted boxes.

observation throughout the oceans of the world. There is limited information available on the time–frequency characteristics and source level of the other, less common, fin whale vocalization types.

Many previous studies of marine mammal vocalizations have been based on observations with a single hydrophone, a small number of widely separated hydrophones or sparse sensor array (Watkins, 1981; Charif *et al.*, 2002), and ocean bottom seismometers (Gaspà Rebull *et al.*, 2006; Harris *et al.*, 2013; Matias and Harris, 2015) to provide detection and classification of the vocalizations, with some localization and tracking of the received vocalizations (Charif *et al.*, 2002). The ability to monitor and differentiate vocalizations from a given marine mammal species can often be challenging when there are multiple marine mammal species vocalizing in close proximity and when the vocalizations are received in overlapping time periods and frequency bands, especially with single hydrophone measurements. Furthermore, it is also challenging to estimate whale ranges from vocalizations received on a single hydrophone or a sparse array. Therefore, hydrophone arrays have been used in the past to locate and track fin whales from their vocalizations (Clark and Fristrup, 1997; Croll *et al.*, 2002; Wang *et al.*, 2016a). The advantage of a coherent hydrophone array, such as the one used here, is that the bearings and times of the received whale vocalizations can be directly estimated and the signal-to-noise ratios (SNRs) enhanced via beamforming. The beamforming enables calls coming from both nearby and distant whales, in different azimuthal bearings relative to the coherent hydrophone array, to be distinguished and separated. Furthermore, long-term monitoring of whale vocalization bearing-time trajectories enables uncommon or previously unobserved calls to be associated or differentiated from the known or commonly observed calls of a given whale species. Locations of whale vocalizations can be readily estimated from their measured bearing-time trajectories (Gong *et al.*, 2013, 2014, 2015) and utilized to generate temporal–spatial distributions of whale vocalizations. POAWRS was

previously applied to detect, localize, and classify the vocalization signals from multiple baleen whale species that include the fin whale (Wang *et al.*, 2016a, b), and toothed whale species simultaneously in the Gulf of Maine (Gong *et al.*, 2014; Huang *et al.*, 2016; Wang *et al.*, 2016a, b), as well as from sperm whales along the US east coast (Tran *et al.*, 2014) over continental shelf-scale regions approximately 100 000 km^2 in size. Temporal–spatial distributions of marine mammal vocalizations from diverse species, based on POAWRS sensing over wide areas, have been overlain with Atlantic herring fish population spatial density distributions in the Gulf of Maine to provide insights into the predator–prey dynamics in that ecosystem (Wang *et al.*, 2016a).

Here the POAWRS technology is applied to observe fin whale vocalizations instantaneously over wide areas spanning 360 degrees in horizontal azimuth and roughly 50–100 km in range, depending on the bathymetry, using a coherent hydrophone array in the Norwegian and Barents Seas. The objectives of this article are to (i) provide a time–frequency characterization for different call types observed (20 Hz pulses, 130 Hz upsweeps, 30–100 Hz downsweep chirps, and 18–19 Hz backbeats); (ii) compare their relative abundances in three different coastal regions off Alesund, Lofoten, and Northern Finnmark; (iii) estimate the temporal and spatial distributions, (iv) source level distributions, and (v) probability of detection (PoD) regions for the more abundant 20 Hz pulse and 130 Hz upsweep call types. This observation and analysis provide insights into the mechanism for sound production in fin whales. The large volume of fin whale 130 Hz upsweep vocalizations observed here are used to investigate whether fin whales are capable of producing this vocalization type independently of their other vocalization types, since they were previously observed in very small quantities and at time instances coinciding with 20 Hz pulses. The findings presented here on fin whale vocalization distribution and behaviour can be applied in future studies of predator–prey interactions in the Norwegian Sea.

Material and methods

Measurement of fin whale vocalizations using a coherent hydrophone array

The underwater recordings of fin whale vocalizations analysed here are drawn from the Norwegian Sea 2014 Experiment (NorEx14), conducted by a collaborative team from the Massachusetts Institute of Technology, Northeastern University, NOAA-Northeast Fisheries Science Center, Naval Research Laboratory, Penn State University, and the Woods Hole Oceanographic Institution in the United States, as well as the Institute of Marine Research-Bergen (IMR) in Norway. The NorEx14 was conducted from 18 February to 8 March 2014, in conjunction with the IMR survey of spawning populations of Atlantic herring off the Alesund coast, the Atlantic cod off the Lofoten peninsula, and the capelin off the Northern Finnmark region. The twofold objectives of the NorEx14 were to (i) image and monitor the population distributions of these large fish shoals from diverse species instantaneously over wide areas of their spawning grounds using the ocean acoustic waveguide remote sensing (OAWRS) and imaging system (Makris *et al.*, 2006, 2009; Jagannathan *et al.*, 2009) from which fish group behavioural patterns could be quantified; and (ii) observe marine mammal vocalizations and infer their temporal–spatial distributions over wide areas using the POAWRS technique (Gong *et al.*, 2014; Huang *et al.*, 2016; Wang *et al.*, 2016a), combined with visual observations for species confirmation. These results would then be employed in future studies of predator–prey interaction and dynamics. The marine mammal vocalization data, that include fin whale vocalizations obtained from POAWRS sensing, were partially processed at sea and further analysed in post-processing.

Table 1. POAWRS receiving array 1-dB angular width $\beta_{1\,dB}(\varphi, f_c)$ at broadside ($\varphi = 0$) and endfire ($\varphi = \pi/2$), given ULF aperture length L, as a function of centre frequency f_c for a given fin whale call type.

Fin whale	f_c	L	$\beta_{1\,dB}\,(\varphi = 0)$	$\beta_{1\,dB}\,(\varphi = \pi/2)$
Call type	(Hz)	(m)	(deg)	(deg)
20 Hz pulse	21.5	189	10	19.5
130 Hz upsweep	128.7	189	1.7	8

The amplitude weighted average frequency values in Table 2 were used as the centre frequency values. A Hanning spatial window is applied in the beamforming.

In NorEx14, recordings of underwater sound were acquired using a horizontal coherent hydrophone array (Becker and Preston, 2003) towed at an average speed of 4 knots (roughly 2 m/s) along designated tracks for 8–24 hours per day. To minimize the effect of tow ship noise on the recorded acoustic data, the coherent hydrophone array was towed approximately 280–330 m behind the research vessel so as to confine this noise to the forward endfire direction of the array, which is the forward direction parallel to the array axis. The tow ship noise in directions away from the forward endfire was negligible after coherent beamforming. The water depth ranged from 100 to 300 m at the array locations, and the array tow depth varied from 45 to 70 m in NorEx14.

The multiple nested sub-apertures of the array contain a total of 160 hydrophones spanning a frequency range from below 15 to 4000 Hz for spatially unaliased sensing. The mean sensitivity of each hydrophone is a constant in this frequency range. A fixed sampling frequency of 8000 Hz was used so that acoustic signals with frequency contents up to 4000 Hz were recorded without temporal aliasing. The ultra low-frequency (ULF) sub-aperture of the array, consisting of 64 equally spaced hydrophones with inter-element spacing of 3 m, was used here to collect fin whale vocalizations with frequency content below 250 Hz. The horizontal beamwidth of the array is a function of the array aperture length L, steering angle φ, as well as centre frequency f_c and bandwidth B of the signal (Johnson and Dudgeon, 1992; Makris *et al.*, 1995; Ratilal *et al.*, 2005). The 1 dB angular width $\beta_{1dB}(\varphi, f_c)$ (Tran *et al.*, 2014) of the receiver array for the fin whale 20 Hz pulse and 130 Hz upsweep vocalizations are provided in Table 1. The steering angle φ is measured as the horizontal azimuthal angle from array broadside. The bearing estimation errors are significantly smaller by a factor of roughly 1/5 for the fin whale 130 Hz upsweep signals in comparison to the 20 Hz pulse signals, as can also be noted in Figure 5, after beamforming with the ULF sub-aperture.

Physical oceanography was monitored by sampling water-column temperature and salinity with expendable bathy thermographs (XBTs) and conductivity–temperature–depth (CTD) sensors at regular intervals of a couple of hours each day. The water-column sound speed profile measured in the three distinct regions of the Norwegian Sea are provided in Jain (2015).

The detection of long-range propagated sounds is significantly enhanced by spatial beamforming and spectrogram analysis which

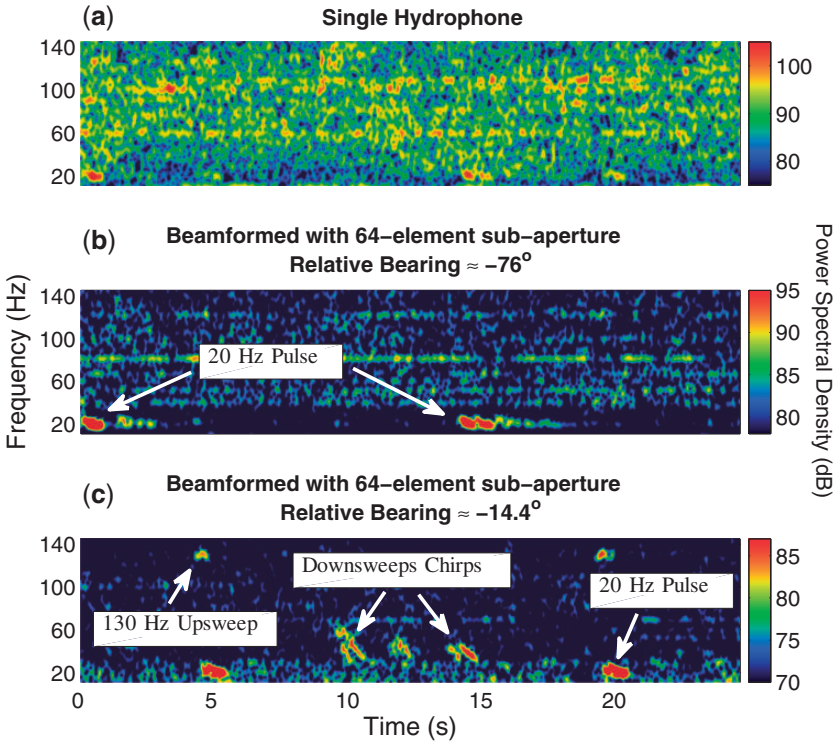

Figure 2. Coherent array processing enhances the SNR to aid in the detection of fin whale vocalizations. Compare single hydrophone measured spectrogram in (a) with spectrogram after coherent beamforming in (b) and (c) with 64-element ULF sub-aperture of POAWRS 160-element hydrophone array. The fin whale vocalizations from the POAWRS receiver array, recorded on 20 February 2014 at 22: 59: 19 GMT, is enhanced by up to 5.3 dB for the 20 Hz pulse and 13.7 dB for the 130 Hz upsweep. In plot (b), two 20 Hz pulses are visually detected above the background noise after beamforming where the fin whale bearing is $\approx -76^{o}$ from array broadside. In plot (c), two 20 Hz pulses, two 130 Hz upsweeps, and three higher frequency downsweep chirps are visually detected above the background noise after beamforming where the fin whale bearing is $\approx 14.4^{o}$ from array broadside.

filters the background noise that is outside of the beam and frequency band of the fin whale vocalizations. The high gain (Johnson and Dudgeon, 1992; Kay, 1998) of the coherent 64-hydrophone ULF sub-aperture, of up to $10 \log_{10} 64 = 18$ dB, enabled detection of fin whale vocalizations up to two orders of magnitude more distant in range in the shallow water environment than a single omnidirectional hydrophone, which has no array gain (Figure 2). The actual array gain, which may be smaller than the full 18-dB theoretical array gain, is dependent on noise coherence and vocalization wavelength relative to array aperture length. For example, the array gain for the 20 Hz pulse is 5.3 dB, while the array gain for the 130 Hz upsweep is 13.7 dB due to the difference in wavelengths of the signals. The array gain is tabulated in Table 5.

The POAWRS coherent hydrophone array employed in NorEx14 detected significant sounds from a wide range of underwater acoustic sources including marine mammal vocalizations from diverse baleen and toothed whale species in the frequency range from 10 Hz up to 4 kHz, and sounds from a large number of diesel-electric surface ships and other powered ocean vehicles (Huang *et al.*, 2017). Here the analysis is focused on the detection and characterization of fin whale vocalizations between the 10 Hz and 200 Hz frequency range. Concurrent ship-based visual observations conducted during our experiment provides confirmation of the presence of fin whales.

Fin whale vocalization detection, bearing estimation, and characterization

Acoustic pressure time series measured by sensors across the receiver array were converted to two-dimensional beam-time series by beamforming (Johnson and Dudgeon, 1992). A total of 64 beams were formed spanning 360 degree horizontal azimuth about the receiver array for data from the ULF sub-aperture. Each beam-time series was converted to a beamformed spectrogram by short-time Fourier transform (sampling frequency = 8000 Hz, frame = 2048 samples, overlap = 3/4, Hann window). Significant sounds present in the beamformed spectrograms were automatically detected by first applying a pixel intensity threshold detector (Sezan, 1990) followed by pixel clustering, and verified by visual inspection (Huang *et al.*, 2016; Wang *et al.*, 2016a, b; Huang *et al.*, 2017). Beamformed spectrogram pixels with local intensity values that are 5.6 dB above the background are grouped using a clustering algorithm according to a nearest-neighbour criteria that determines if the pixels can be grouped into one or more significant sound signals. Each individual detected signal is next characterized by its pitch track (Wang and Seneff, 2000; Shapiro and Wang, 2009; Baumgartner and Mussoline, 2011) representing the time variation of the fundamental frequencies. The pitch-track is estimated using a time-frequency peak detector

from a signal's detected and clustered pixel intensity values in the beamformed spectrogram.

The horizontal azimuthal direction or bearing $\hat{\varphi}$ of each detected signal, measured from array broadside, is estimated using a beamforming technique (Johnson and Dudgeon, 1992) that selects the bearing in which the beamformed, band-pass filtered pressure–time series contained maximum energy during the time duration of the signal and in the same frequency band. The estimated relative bearings $\hat{\varphi}$, measured with respect to array broadside, are then converted to absolute bearings, measured from the array centre with respect to true North.

The time-–frequency characteristics of each individual detected signal is determined from its pitch-track. The pitch-track for a signal contains a time series $t = (t_1, t_2, \ldots, t_i)$, a frequency series $f = (f_1, f_2, \ldots, f_i)$, and an amplitude series $A = (A_1, A_2, \ldots, A_i)$ describing the time variation of the fundamental frequency in the signal (Wang and Seneff, 2000; Shapiro and Wang, 2009; Baumgartner and Mussoline, 2011). Eight features are extracted from each signal. They are (1) minimum frequency (Hz), f_L; (2) maximum frequency (Hz), f_U; (3) amplitude weighted average frequency (Hz), $\bar{f}$; (4) mean instantaneous bandwidth (Hz), $\bar{B}$; (5) relative instantaneous bandwidth, $\bar{B}/\bar{f}$; (6) duration (s), $\tau = t_i - t_1$; (7) slope from first-order polynomial fit (Hz/s), $\frac{df}{d\tau}$; and (8) curvature from second-order polynomial fit (Hz/s^2), $\frac{d^2f}{d\tau^2}$. The slope and curvature are obtained from second-order nonlinear curve-fit to the vocalization traces obtained via pitch-tracking (Huang *et al.*, 2016; Wang *et al.*, 2016a).

The time–frequency characteristics extracted via pitch tracking are applied for fin whale vocalization classification. A combination of extracted features from pitch-tracking, orthogonalized via principle component analysis (PCA) (Jolliffe, 2002), were used to optimize the vocalization classification employing *k*-means (Kanungo *et al.*, 2002) and Bayesian-based Gaussian mixture model clustering approaches (Richard *et al.*, 2001). The number of clusters can be determined via Bayesian information criterion (Hirose *et al.*, 2011). The bearing-time trajectories of each closely associated series of vocalizations were also taken into account to ensure consistent classification (Huang *et al.*, 2016).

Determination of detected fin whale vocalization rate and time series over the diel cycle

The diel vocalization rate in units of calls/day and vocalization rate time series in units of calls/min for the detected fin whale 20 Hz pulses and 130 Hz upsweeps were obtained by averaging the vocalization rate time series for that type of vocalization over multiple diel cycles in a specific region (Figure 6). The detected fin whale vocalization rate time series are averaged over 15 min bins.

Localization of fin whale vocalizations

The horizontal location of each detected fin whale vocalization consists of a range and a bearing estimate. The estimated azimuthal bearings of sequences of fin whale vocalizations form multiple bearing-time trajectories (Figure 5). These bearing-time trajectories are utilized to determine the ranges of the fin whale vocalizations from the horizontal receiver array centre employing the moving array triangulation (MAT) (Gong *et al.*, 2013, 2014, 2015) and the bearings-migration minimum mean-square-error (MMSE) methods (Gong *et al.*, 2013). Position estimation error, or the root-mean-square (RMS) distance between the actual and estimated location, is a combination of range and bearing errors.

The bearing estimation error of the time-domain beamformer for the 130 Hz upsweep and the 20 Hz pulse vocalizations were provided in Table 1. These bearing accuracies for our beamformer have been verified by both theoretical modelling (see formulation and numerical implementation in Figures 3 and 4 of Wang and Ratilal, 2017) and application to experimental data (see Figure 6 of Tran *et al.*, 2014 showing beamformer output and resolution for broadside and endfire arrivals of a broadband signal).

The range estimation errors have been quantified for this array (Gong *et al.*, 2013, 2014, 2015) for broadband signals with roughly 5% bandwidth to centre frequency ratios, and with approximately 50% centre frequency to array aperture design frequency ratios. The fin whale 130 Hz upsweep vocalizations fall within this category of broadband signals. For such signals, the range estimation error, expressed as the percentage of the range from the source location to the horizontal receiver array centre, for the MAT and MMSE is roughly 2% at array broadside and gradually increases to 10% at 65° from broadside and 25% near or at endfire. These errors are determined previously from thousands of controlled source signals transmitted by a source array, and are based on absolute global positioning system (GPS) ground truth measurements of the source array's position (Gong *et al.*, 2013, 2015). Note that the range estimation error for the fin whale 20 Hz pulses are expected to be larger than that for the 130 Hz upsweeps, because of the larger bearing estimation error at 20 Hz.

More than 85% of fin whale vocalizations are found to be located from 0° to 65° from the broadside direction of the horizontal hydrophone array. Position estimation error is less than 2 km and 5 km, respectively, for majority of the fin 130 Hz and 20 Hz vocalizations localized since they are found mostly within 50 km of the horizontal receiver array centre. This error is roughly an order of magnitude smaller than or equivalent to the spatial scales of the fin whale concentrations shown in Figures 12 and 13, and consequently has negligible influence on the analyses and results.

Detected fin whale vocalization rate spatial density distributions

The estimated locations for detected fin whale vocalizations over the duration of our data collection are used to generate the fin whale vocalization rate spatial density distribution maps shown in Figures 12 and 13. The location of each fin whale vocalization is characterized by a 2 D Gaussian probability density function with mean equal to the measured mean position from MAT and standard deviation ellipse with major and minor axes determined by the measured range and bearing standard deviations. The detected fin whale call rate spatial density distribution map for a specific fin whale call type is determined by superposition of the 2 D spatial probability densities for the location of each call, normalized by the total measurement time. This approach for estimating the detected vocalization rate spatial density distribution was previously applied to fin whales and other baleen whale species in Gong *et al.* (2014), Huang *et al.* (2016), and Wang *et al.* (2016a).

Note that the detected vocalization rate spatial density distributions calculated here quantify the mean fin whale call volume within time units of 1 min and within areal units of 25 nmi^2 averaged over multiple diel cycles. The translation of a fin whale individual at swim speeds ranging from 0 to 9 m/s over a minute time

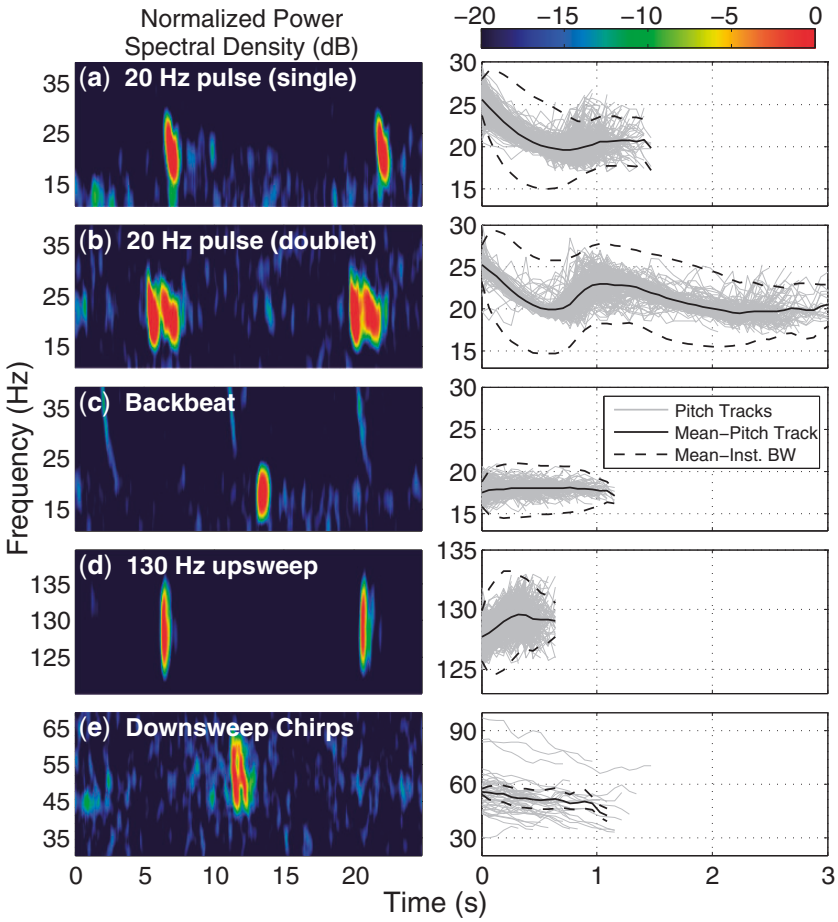

Figure 3. Spectrograms and pitch-tracks for the fin whale (a) 20 Hz pulse (single), (b) 20 Hz pulse (doublet), (c) backbeat, (d) 130 Hz upsweep, and (e) 30–100 Hz downsweep chirp. The pitch-track figures are displayed utilizing (a) 963, (b) 218, (c) 289, (d) 828, and (d) 48 fin whale vocalizations. The mean pitch-track is indicated by the solid black curve, while the mean instantaneous bandwidth of the pitch-tracks are indicated by the dashed black curves.

interval is bounded by 0.54 km which is significantly smaller than the spatial extent of each 25 nmi^2 areal unit that the data are averaged over, and therefore has negligible effect on the detected fin whale vocalization rate spatial density distribution maps shown in Figures 12 and 13. Furthermore, a vocalizing fin whale travelling at 0–9 m/s speed would require more than 18 min to traverse across the 25 nmi^2 areal unit. During this time frame, the MAT technique provides numerous independent estimates of the vocalizing fin whale horizontal position, so that the errors in mean position estimate and tracking estimates can be reduced by statistical averaging (Makris, 1996; Frieden, 2012; Tran *et al.*, 2012).

The effects of detector performance on the POAWRS determined vocalization rate spatial density distribution have been previously investigated, where the PoD-normalized distributions were compared with the distributions without the PoD normalization for fin whale and other marine mammal species in the Gulf of Maine (refer to Section III of the Supplementary Information of Wang *et al.*, 2016a and compare Supplementary Information Figures 2 and 5 where the PoD-normalized distribution is compared with those without PoD normalization). The detector performance has negligible effect on the vocalization rate spatial density distribution in regions where the PoD is high $> 80\%$. In Figures 12 and 13, the distribution shown is valid in the region bounded by the dashed lines where the PoD is high for both the fin whale 20 and 130 Hz vocalizations.

Source level estimation for fin whale vocalizations

The fin whale vocalization source level $L_S(\mathbf{r}_0)$ is estimated (Figures 7 and 8) using the passive sonar equation (Urick, 1983; Kinsler *et al.*, 1999; Gong *et al.*, 2014):

$$L_S(\mathbf{r}_0) = \mathrm{RL}(\mathbf{r}) + \mathrm{TL}(|\mathbf{r} - \mathbf{r}_0|), \quad (1)$$

where $\mathrm{RL}(\mathbf{r})$ is the received whale vocalization pressure level for a receiver located at $\mathbf{r}$. The received whale vocalization pressure level was estimated as the RMS value of the maximum instantaneous time-domain signal bandpass-filtered between upper f_U and lower f_L frequencies and beamformed to the azimuthal bearing of the vocalization, over a time window (Madsen and Wahlberg, 2007) encompassing 90% of the total signal energy. The upper f_U and lower f_L frequencies are determined as the -10 dB end points relative to the signal peak in the power spectrum.

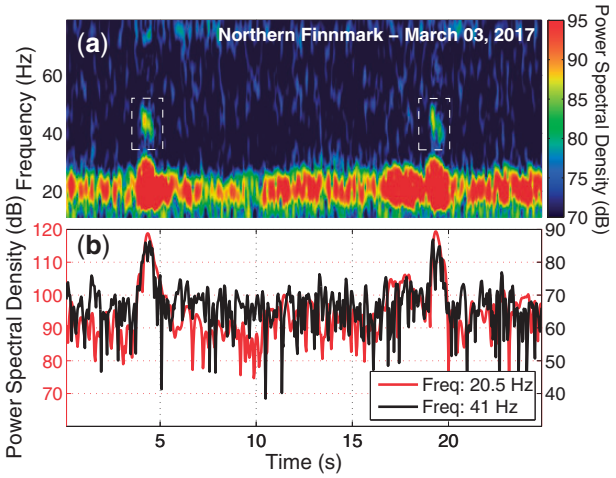

Figure 4. Example of fin whale 20 Hz pulse vocalizations also containing the second harmonic at twice the peak centre frequency, and roughly 25–30 dB lower received pressure level, measured by the coherent hydrophone array from a nearby fin whale during NorEx14. Plot (a) displays two 20 Hz pulses containing the second-order harmonic (highlighted in dashed white box), while plot (b) displays the power spectral density versus time for the 20 Hz pulse peak centre frequency (centred around 20.5 Hz) and the second-order harmonic peak centre frequency (centred around 41 Hz).

The corresponding one-way broadband acoustic transmission loss, $\mathrm{TL}(|\mathbf{r}-\mathbf{r}_0|)$, from the estimated location of each fin whale vocalization to the centre of the POAWRS receiver array, was calculated using a calibrated (Jain, 2015; Schory, 2015) parabolic equation-based range-dependent acoustic propagation model (RAM) (Collins, 1993):

$$\mathrm{TL}(|\mathbf{r}-\mathbf{r}_0|) = 10\,\log_{10}\Big(\int_{f_L}^{f_U} Q(f)\langle|G(\mathbf{r}|\mathbf{r}_0,f)|^2\rangle \mathrm{d}f\Big), \qquad (2)$$

where $G(\mathbf{r}|\mathbf{r}_0,f)$ is the waveguide Green function at frequency f for a whale located at $\mathbf{r}_0$ and receiver at $\mathbf{r}$, $Q(f)$ is the normalized vocalization spectra, and f_U and f_L are the upper and lower frequencies used for the bandpass filter. The model takes into account the environmental parameters such as the range-dependent water depth and sound speed profiles measured in the Norwegian and Barents Seas to stochastically compute the propagated acoustic intensities via Monte-Carlo simulations following the approach of Andrews *et al.* (2009), Gong *et al.* (2010), and Andrews *et al.* (2011). The mean magnitude-squared waveguide Green function is obtained by averaging over multiple Monte-Carlo simulations, weighted by the whale call depth probability density function distribution from the sea surface to the sea floor to account for waveguide fluctuations and the unknown whale depth, respectively. Here the whale call depth probability density function is modelled as a Gaussian random variable with a mean of 13 m and a standard deviation of 6 m [following approximately the findings presented in Figure 4a of Stimpert *et al.*, 2015]. The broadband transmission loss standard deviations are calculated in the log-transformed domain using the broadband transmission loss at each potential whale depth from the sea surface to the seafloor.

The approaches for estimating the vocalization source level and modelling the broadband transmission loss were previously applied in the Gulf of Maine for fin whales and several other baleen whale species (Gong *et al.*, 2014; Wang *et al.*, 2016a, b).

Probability of detection regions for fin whale vocalizations

The POAWRS PoD $P_D(r)$ for a specific fin whale vocalization, as a function of range r from the coherent hydrophone array, is modelled using the approach provided in Appendix 1. We model the PoD regions for the fin whale 20 Hz pulse and 130 Hz upsweep vocalizations received on a coherent hydrophone array after spatial beamforming. The fin whale vocalization source levels L_s used in the PoD calculations are estimated from a subset of the POAWRS received vocalizations, in units of dB re 1 μPa at 1 m, and is determined using the results in Figures 7 and 8. We model the PoD regions off the coast of Alesund (region I), Lofoten (region II), and Northern Finnmark (region III) (Figure 1). This approach was previously applied to estimate the PoD regions for fin 20 Hz vocalizations in the Gulf of Maine [see Supplementary Information section I and Supplementary Information Figure 1(b) of Wang *et al.*, 2016a, showing the 10%, 30%, 50%, 70%, and 90% PoD regions specifically for fin whale vocalizations received on a coherent hydrophone array].

Results

Here we first identify and describe the repertoire of fin whale vocalizations, in the 10–200 Hz frequency range comprising of a variety of call types, measured by the coherent hydrophone array during the NorEx14. We next provide a statistical time–frequency characterization of each fin whale vocalization type observed. Typical examples of measured fin whale vocalization bearing-time trajectories are provided for the Alesund, Lofoten, and Northern Finnmark offshore regions. We determine the mean diel call volumes, the diel vocalization rate time series, and the source level distributions for the two most prominent fin whale vocalization types measured in NorEx14. The estimated source level distributions are then applied to calculate the PoD regions for the corresponding fin whale vocalization types for both the coherent hydrophone array and for a single hydrophone. Finally, we provide examples of detected fin whale vocalization rate spatial density distributions based on diel monitoring in the Alesund, Lofoten, and Northern Finnmark coastal regions.

Fin whale vocalization types and time–frequency characterization

During the NorEx14, the main types of fin whale vocalizations observed were the 20 Hz pulse, the 18–19 Hz backbeat pulse, the 130 Hz upsweep pulse, and the 30–100 Hz downsweep chirp (Watkins *et al.*, 1987; Clark and Gagnon, 2004; Simon *et al.*, 2010; Castellote *et al.*, 2012). It should be noted that a subset of the 30–100 Hz downsweeps detected during the NorEx14 are in the same frequency range as the fin whale 40 Hz call identified by Širović *et al.*, 2013. It seems highly likely that the 40 Hz call may just be a subset of the 30–100 Hz downsweeps identified by Watkins *et al.* (1987) and Castellote *et al.* (2012). The 20 Hz pulses were observed in two specific patterned sequences which are a repeated sequence of one 20 Hz pulse and a repeated sequence of two consecutive 20 Hz pulses. In this article, we will refer to these two specific patterned sequences as the 20 Hz pulse (single) and the 20 Hz pulse (doublet) (Watkins *et al.*, 1987; Thompson *et al.*, 1992; Croll *et al.*, 2002). These two fin whale

Table 2. Estimated pitch-track features for various types of fin whale vocalizations observed during NorEx14.

Characteristics	20 Hz pulse (single)	20 Hz pulse (doublet)	130 Hz pulse upsweep	Downsweep chirps	Backbeat
***n* (no. analysed)**	**963**	**368**	**1664**	**39**	**289**
f_L (Hz)	14.7±0.8	14.5±0.8	124.0±1.3	43.0±9.3	14.0±0.7
f_U (Hz)	29.2±0.9	30.0±1.1	133.9±1.4	61.0±11.7	21.7±1.3
$\bar{f}$ (Hz)	21.5±0.5	21.5±0.4	128.7±0.7	51.4±10.1	17.9±0.5
$\bar{B}$ (Hz)	9.0±0.8	9.5±0.9	6.6±1.4	8.4±3.3	5.4±1.4
$\bar{B}/\bar{f}$	0.43±0.04	0.44±0.04	0.051±0.011	0.17±0.07	0.30±0.08
τ (s)	0.88±0.18	2.5±0.3	0.42±0.15	0.69±0.28	0.69±0.21
$\frac{df}{d\tau}$ (Hz/s)	−6.7±2.3	−1.9±0.4	4.3±5.1	−12.2.±10.5	0.3±1.7
$\frac{d^2f}{d\tau^2}$ (Hz/s^2)	11.0±5.4	−0.22±1.3	−7.5±36.0	0.9±37.7	−5.0±12.4

vocalization sequences have been previously documented in Watkins (1981). Typical spectrograms, as well as mean and ensemble pitch-tracks for a subset from each fin whale vocalization type observed, are shown in Figure 3. The time–frequency characteristics estimated from pitch-tracking and the IPIs of the repetitive fin whale vocalization types are tabulated in Tables 2 and 3, respectively. The subset of vocalizations analysed in the tables were received with significantly high SNR (SNR $>$ 10 dB).

The time–frequency characteristics of each fin whale vocalization extracted from pitch-tracking are applied to classify the vocalization according to fin whale call type (Table 2). The fin whale 20 Hz pulse (single) and the 20 Hz pulse (doublet) vocalizations share several similar frequency characteristics since both vocalization sequences are based on the 20 Hz pulse. The time duration, however, is a good discriminant between the two vocalization types. The time duration of the 20 Hz pulse (single) has a mean and standard deviation of 0.88 ± 0.18 s, while that for the 20 Hz pulse (doublet) is 2.5 ± 0.3 s, which is roughly 2.5 times larger. Two potentially strong discriminants between the 20 Hz pulse (single) and the backbeat vocalizations are the maximum frequency (f_U) and slope ($df/d\tau$), which are 29.2 ± 0.9 Hz and −6.7±2.3 Hz/s, respectively for the 20 Hz pulse (single), while they are 21.7 ± 1.3 Hz and 0.3 ± 1.7 Hz/s for the backbeat vocalizations. The pitch-tracks of the backbeat vocalizations have on average a positive slope versus time, while the pitch-tracks of the 20 Hz pulse (single) have on average negative slopes (Figure 3). The amplitude-weighted average frequency ($\bar{f}$) of the 130 Hz upsweeps and the 30–100 Hz downsweeps provide a good discriminant between the other fin whale vocalization types. The distributions for $\bar{f}$ are characterized by a mean and standard deviation of 128.7 ± 0.7 Hz for the 130 Hz upsweeps, and 51.4 ± 10.1 Hz for the 30–100 Hz downsweeps.

An IPI is quantified for patterned sequences of the fin whale 20 Hz pulse (single), 20 Hz pulse (doublet), 130 Hz upsweep, and backbeat pulse (Table 3). No IPI could be quantified for the 30–100 Hz downsweeps since they occurred randomly in time during the NorEx14. The 130 Hz upsweep was observed to occur in combination with the 20 Hz pulse (single) and 20 Hz pulse (doublet) throughout most of the NorEx14. Therefore, the observed 130 Hz upsweeps IPIs matched either a 20 Hz pulse (single), a 20 Hz (doublet), or both. The IPI distributions for the 20 Hz pulse (single), 20 Hz pulse (doublet), and 130 Hz upsweep are characterized by approximately the same mean and standard deviation of 14.5 ± 0.3 s. However, there were two specific fin whale bearing-time trajectories observed with different IPI distributions, which

Table 3. IPIs observed for fin whale vocalizations during NorEx14.

	20 Hz pulse (single)	20 Hz pulse (doublet)	130 Hz pulse upsweep	Downsweep chirps	Backbeat
IPI	14.5±0.3	14.5±0.3	14.5±0.3	None observed	24.7 ±0.4
(*s*)	15.3±0.3	29.7±0.4	15.3±0.3		
			29.7±0.4		

are characterized by a mean and standard deviation of 15.3 ± 0.3 s on 23 February 2014 (Alesund region), and 29.7 ± 0.3 s on 26 February 2014 (Northern Finnmark region). Each of these two distributions are associated with only one specific fin whale bearing-time trajectory during the day of observation. The IPI distribution for the backbeat pulse is calculated from observations on 20 February 2014 (Alesund region) and is characterized by a mean and standard deviation of 24.7 ± 0.4 s. This specific fin whale bearing-time trajectory was observed to exclusively contain fin whale backbeat vocalizations with a duration of approximately 25 min for the pulse train. All other observations of the backbeat vocalizations during the NorEx14 occurred with corresponding 20 Hz pulses in the sequences.

Second-order harmonic of the fin whale 20 Hz pulse

For fin whales in close proximity to the coherent hydrophone array, the dominant and intense 20 Hz pulses were received along with their second-order harmonic. It was observed that the received pressure level for the 20 Hz pulse had its peak energy centred around 20.5 Hz, while the second-order harmonic was centred around 41 Hz and roughly 25–30 dB lower received pressure level (Figure 4). This second-order harmonic was observed off Northern Finnmark on 28 February 2014 and 3–4 March 2014. The received pressure level difference between the 20 Hz pulse and corresponding second-order harmonic was estimated using 33 detections on 3 February 2014 with mean of 26.18 ± 2.6 dB, 51 detections on 3 March 2014 with mean of 30.05 ± 2.46 dB, and 48 detections on 4 March 2014 with mean of 26.23 ± 8.71 dB. Harmonics higher than second order were not detected in the beamformed spectrograms.

Detected fin whale vocalization rate and time series over the diel cycle

Fin whale vocalizations were measured by the coherent hydrophone array at each passive acoustic monitoring location off the

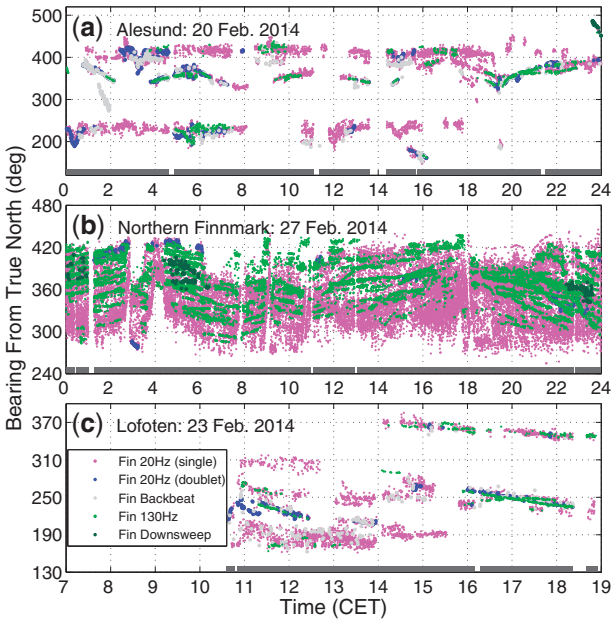

Figure 5. Bearing-time trajectories of fin whale vocalizations detected by the POAWRS coherent hydrophone array on (a) 20 February 2014 (Alesund), (b) 27 February 2014 (Northern Finnmark), and (c) 23 February 2014 (Lofoten).

Alesund (region I), Northern Finnmark (region III), and Lofoten (regions II and IV) coastal regions during the NorEx14 (Figure 1). Typical examples of measured bearing-time trajectories containing fin whale vocalizations in the three regions are shown in Figure 5. The bearing-time trajectories are plotted between 125 and 500 degrees from true north relative to the coherent hydrophone array centre at each time instance in order to make the bearing-time trajectories continuous within a 360 degree azimuth span. In Figure 5, each of the fin whale vocalization types, identified in Figure 3, were detected except for the backbeat pulse on 27 February 2014 (Northern Finnmark region). The most frequently observed fin whale vocalization types in all regions were the 20 Hz pulse and 130 Hz upsweep. The detected fin whale vocalization rates in units of calls/day for each of these call types are estimated by averaging over multiple diel cycles in each observation region (Table 4). The detected fin whale vocalization rates are found to be significantly higher off the Northern Finnmark coast at roughly 37000 ± 5000 calls/day for the 20 Hz pulse and 21000 ± 5300 calls/day for the 130 Hz upsweep. These vocalization rates are a factor of roughly 5 times and 17 times larger for the 20 Hz pulse and the 130 Hz upsweep, respectively in the Northern Finnmark coastal region than those off the Alesund and Lofoten coasts. The 130 Hz upsweeps comprise roughly 35% of the measured fin whale vocalizations in the Northern Finnmark coastal region. Only roughly 15% of the fin whale vocalizations are comprised of the 130 Hz upsweeps in the Alesund and Lofoten coastal regions.

The detected fin whale vocalization rate time series for the 20 Hz pulse and 130 Hz upsweep are shown in Figure 6. Similar to the diel vocalization rates, the fin whale vocalization rate time series are significantly larger off the Northern Finnmark coast at roughly 20–30 calls/min for the 20 Hz pulse and roughly 10–20 call/min for the 130 Hz upsweep throughout the 24 h time period (Figure 6). These vocalization rate time series are a factor of 2–3 times smaller for the 20 Hz pulses and 10–20 times smaller for the 130 Hz upsweeps off the Alesund and Lofoten coasts.

Fin whale source level distribution estimates for the 20 Hz pulse and 130 Hz upsweep

The source level distributions of fin whale 20 Hz pulses and 130 Hz upsweeps are estimated from vocalizations received by the coherent hydrophone array off the coasts of Alesund (region I), Northern Finnmark (region III), and Lofoten (regions II and IV) during NorEx14 (Figure 1). A subset of measured vocalizations with significantly high SNR (SNR $>$ 10 dB) are employed in the analysis. The mean and standard deviation of the source level distribution, in units of dB re 1 μPa at 1 m, for the 20 Hz pulse is 192.3 ± 6.5 dB off Alesund region, 195.8±4.4 dB off Northern Finnmark region, 186.3 ± 7.1 dB off Lofoten region, and 190.5 ± 7.4 dB by intensity averaging of the results obtained off all three regions (Figure 7). The mean and standard deviation of the source level in units of dB re 1 μPa at 1 m for the 130 Hz upsweep pulse is 170.5 ± 6.0 dB off Alesund region, 171.2 ± 5.0 dB off Northern Finnmark region, 168.5 ± 3.5 dB off Lofoten region, and 170.3±5.2 dB by intensity averaging of the results obtained off all three regions (Figure 8). The mean source level of the fin whale 130 Hz upsweep is approximately 20 dB lower than that of their 20 Hz pulse.

Fin whale probability of detection (PoD) regions for the 20 Hz pulse and 130 Hz upsweep

The PoD regions of the fin whale 20 Hz pulse and 130 Hz upsweep vocalizations received by the coherent hydrophone array in the Alesund (region I), Northern Finnmark (region III), and Lofoten (region II) offshore regions are shown in Figures 9–11. The source level distributions of the fin whale vocalizations used in the PoD calculations for all regions are based on the location averaged result of 190.5 dB re 1 μPa at 1 m for the 20 Hz pulse and 170.3 dB re 1 μPa at 1 m for the 130 Hz upsweep. The array gain (Urick, 1983) values from Table 5 for the two fin whale vocalization types were used in the PoD calculations to account for SNR enhancement after beamforming with the 64-element ULF sub-aperture of the coherent hydrophone array.

The 50% PoD regions for fin whale vocalizations, with the 64-element ULF sub-aperture of the coherent hydrophone array as the receiver, extends over a region more than 200 km in diameter after beamforming (Figures 9–11). In contrast the 50% PoD regions for fin whale vocalizations, with a single omnidirectional hydrophone as the receiver, are comparatively smaller by a factor of 1.5–3 times for the fin whale 20 Hz pulse and by a factor of 4–9 times for the fin whale 130 Hz upsweep.

Despite the 20 dB lower source level of the fin whale 130 Hz upsweep in comparison to that of their 20 Hz pulse, the 50% PoD regions for both these fin whale vocalization types are roughly equivalent when the coherent hydrophone array is employed as the receiver. This is because the array provides significantly higher gain after beamforming for the fin whale 130 Hz upsweeps (13.7 dB array gain) than the 20 Hz pulses (5.3 dB array gain). In contrast the single omnidirectional hydrophone provides no array gain and so the 50% PoD region is significantly smaller for the fin whale 130 Hz upsweep than the 20 Hz pulse because of the 20 dB lower source level of the 130 Hz upsweep.

Table 4. Detected fin whale diel vocalization rates (calls/day) based on measurements with the POAWRS coherent hydrophone array during NorEx14.

Location (coastal)	Audio Recording (h)	20 Hz pulse (calls/day)	130 Hz pulse upsweep (calls/day)	Downsweep chirps (calls/day)	Backbeat (calls/day)
Alesund	16.9 ±5.0	7211 ±3825	1199 ±783	14 ±16	341 ±212
Finnmark	14.4 ±5.5	37290 ±5014	20775 ±5300	*not accessible	*not accessible
Lofoten	4.2 ±2.5	6803 ±5474	1199 ±1027	10 ±26	291 ±370

The results are averaged over multiple diel cycles for each region: 18–21 February 2014 off Alesund (region I), 23 February (region II) and 5–8 March 2014 (region IV) off Lofoten, 26 February–1 March 2014 and 3–4 March 2014 (region III) off Northern Finnmark. (*The diel vocalization rates could not be confidently estimated for the downsweep chirps and backbeats measured off the coast of the Northern Finnmark region due to multiple known and unknown marine mammal species vocalizing in close proximity and in the same frequency band).

Table 5. Parameters used in modelling the PoD regions for fin whale vocalizations in the Norwegian Sea.

Fin whale Call type	$\bar{f}$ (Hz)	$\bar{B}$ (Hz)	L_S (dB re 1 μPa at 1 m)	$S_N(f)$ (dB re 1 μPa/Hz)	AG
20 Hz pulse	21.5	9.0	190.5	93.2	5.3
130 Hz upsweep	128.7	6.6	170.3	84.8	13.7

They are the amplitude-weighted average frequency, $\bar{f}$ (Hz); mean instantaneous bandwidth, $\bar{B}$ (Hz); source level, L_S (dB re 1 μPa at 1 m); omnidirectional ambient noise spectral density level, $S_N(f)$ (dB re 1 μPa/Hz); and coherent beamforming gain of the passive receiver array, AG.

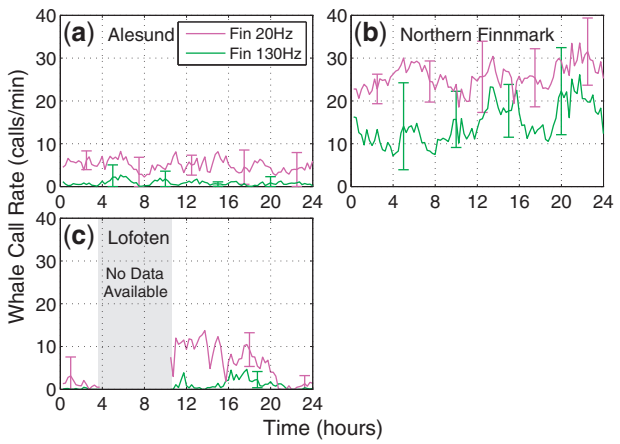

Figure 6. Mean detected fin whale vocalization rate time series (calls/min) for the 20 Hz pulse and 130 Hz upsweep detected off the coasts of (a) Alesund, (b) Northern Finnmark, and (c) Lofoten during NorEx14. The error bars indicate standard deviations obtained from averaging the time series over multiple diel cycles in (a) Alesund from 18 to 21 February 2014, in (b) Northern Finnmark from 26 February 2014 to 4 March 2014, and in (c) Lofoten near Røst on February 23, 2014 and Lofoten near Andenes from 5 to 8 March 2014.

Location-dependent vocalization rate spatial density distributions

The vocalization rate spatial density distributions for the fin whale 20 Hz pulse and 130 Hz upsweep detected off the Alesund region on 20 February 2014, the Lofoten region on 23 February 2014, and the Northern Finnmark region on 27 February 2014

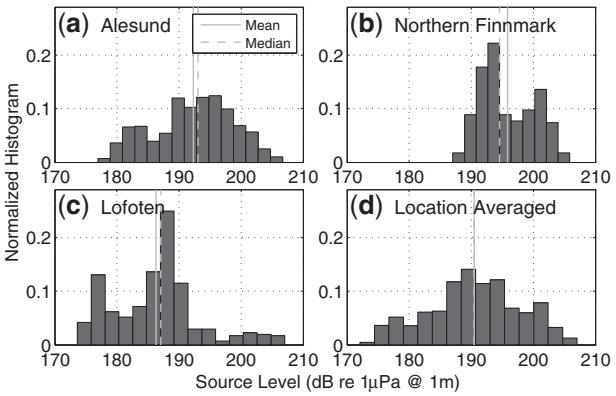

Figure 7. Normalized histograms of the estimated source levels for the fin whale 20 Hz pulses measured in the coastal regions off (a) Alesund, (b) Northern Finnmark, (c) Lofoten, and (d) the combined results from (a), (b), and (c) during the NorEx14. The histograms in (a), (b), (c), and (d) were generated using 724, 691, 323, and 1738 independent estimates of the instantaneous source level.

(Figures 12 and 13) are estimated from localization of the measured bearing-time trajectories of these vocalization in Figure 5. In the Northern Finnmark offshore region, the volume of both the fin whale 20 Hz pulse and 130 Hz upsweep were substantial and span the entire northern hemisphere roughly ±60 degrees from true north.

Discussion

Fin whale location-dependent vocalizations, detected call rates and characteristics

Significantly higher fin whale diel vocalization rates were detected by the coherent hydrophone array in the Northern Finnmark offshore region when compared with the Alesund and Lofoten offshore regions during NorEx14. In particular, the fin whale 130 Hz upsweeps were found to be 17 times more abundant and the 20 Hz pulses were 5 times more abundant off Northern Finnmark. This is likely due to fin whale fish feeding activities off Northern Finnmark since their fish prey, the capelin, is highly abundant in this region as the observation time period of NorEx14 coincided with the capelin spawning season.

Statistical time–frequency characterization of fin whale vocalization types, presented in Table 2, are essential for automatic classification and identification of fin whales with passive acoustics. The estimated mean and standard deviation of the time–

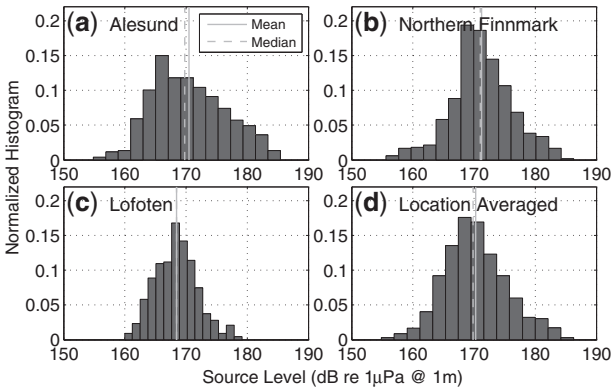

Figure 8. Normalized histograms of the estimated source levels for the fin whale 130 Hz upsweeps measured in the coastal regions off (a) Alesund, (b) Northern Finnmark, (c) Lofoten, and (d) the combined results from (a), (b), and (c) during the NorEx14. The histograms in (a), (b), (c), and (d) were generated using 441, 537, 669, and 1647 independent estimates of the instantaneous source level.

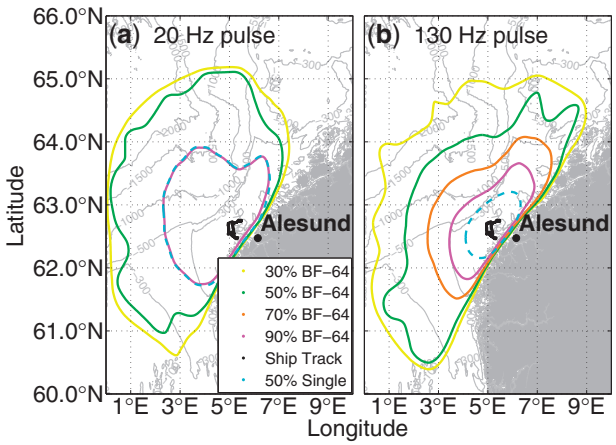

Figure 9. The PoD regions off the Alesund coast for the fin whale (a) 20 Hz pulse and (b) 130 Hz upsweep. The tow tracks of the coherent hydrophone array during 18–21 February 2014 are indicated in solid black. The solid coloured contours provide the % PoD regions for vocalizations received on the 64-element sub-aperture of the coherent hydrophone array after beamforming, while the dashed cyan contour represent the 50% PoD region for vocalizations received on a single omnidirectional hydrophone.

frequency parameters for different fin whale vocalization types can be incorporated into a classifier, such as logistic regression, support vector machine, or decision tree (Richard *et al.*, 2001), to provide automatic classification of the vocalizations according to type, as well as to distinguish fin whale vocalizations from those of other baleen whale species.

Source level estimates for fin whale 20 Hz pulse and 130 Hz upsweep

The fin whale 20 Hz pulse vocalization source level estimates obtained here for the Norwegian Sea compare well with previous estimates for other ocean areas, including the western Antarctic Peninsula (Širović *et al.*, 2007) and Northeast Pacific Ocean

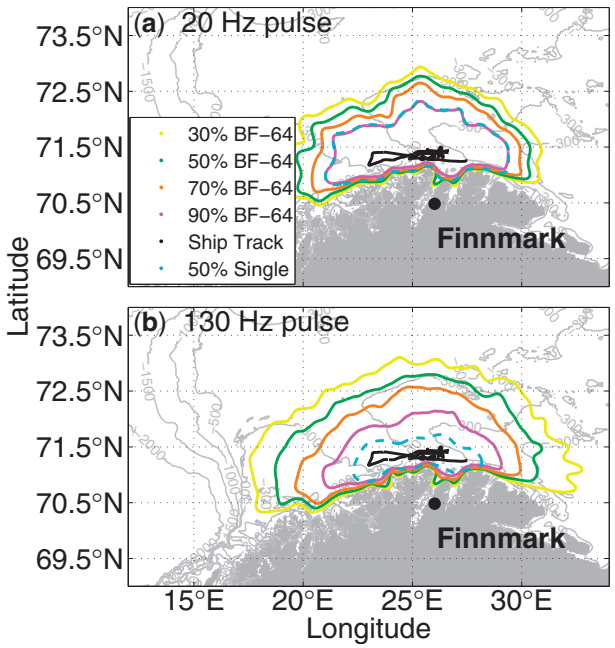

Figure 10. The PoD region off the Northern Finnmark coast for the fin whale (a) 20 Hz pulse and (b) 130 Hz upsweep. The tow tracks of the coherent hydrophone array during 26 February–4 March 2014 are indicated in solid black. The solid coloured contours provide the % PoD regions for vocalizations received on the 64-element sub-aperture of the coherent hydrophone array after beamforming, while the dashed cyan contour represent the 50% PoD region for vocalizations received on a single omnidirectional hydrophone.

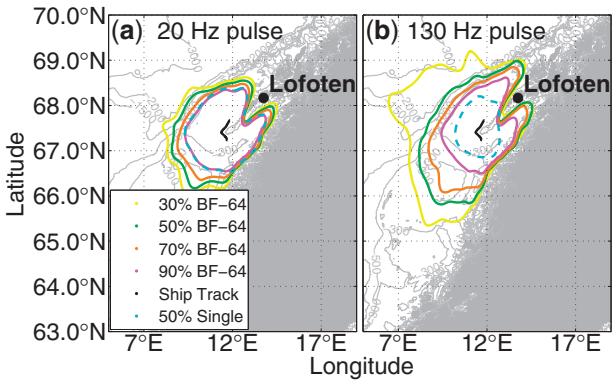

Figure 11. The PoD region off the Lofoten coast (near Røst) for the fin whale (a) 20 Hz pulse and (b) 130 Hz upsweep. The tow tracks of the coherent hydrophone array on 23 February 2014 are indicated in solid black. The solid coloured contours provide the % PoD regions for vocalizations received on the 64-element sub-aperture of the coherent hydrophone array after beamforming, while the dashed cyan contour represent the 50% PoD region for vocalizations received on a single omnidirectional hydrophone.

(Weirathmueller *et al.*, 2013). In general, the range of fin whale vocalization source level estimates from previous studies either overlap well with (Širović *et al.*, 2007; Weirathmueller *et al.*, 2013) or lie fully (Watkins *et al.*, 1987; Wang *et al.*, 2016b) within the range of fin whale vocalization source level estimates obtained here and shown in Figure 7.

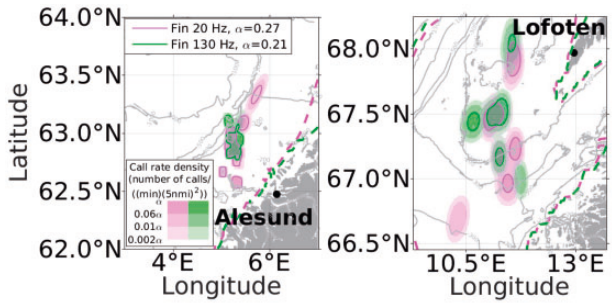

Figure 12. Vocalization rate spatial density distribution maps for the fin whale 20 Hz pulse and 130 Hz upsweep detected in the (a) Alesund coastal region on 20 February 2014, and (b) Lofoten coastal region (near Røst) on 23 February 2014. The fin whale call rate spatial densities in units of number of calls per minute per 25 nmi^2 [(min) (5 nmi)^2] measured by POAWRS have peak values α indicated. The dashed contours represent the 50% PoD regions for the fin whale 20 Hz pulse (purple) and 130 Hz upsweep (green).

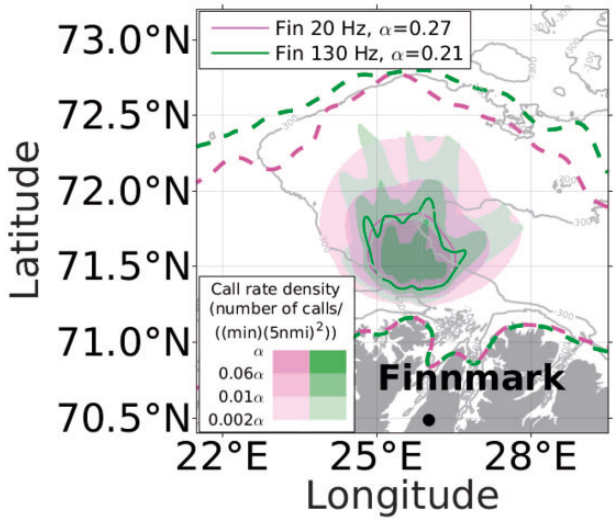

Figure 13. Vocalization rate spatial density distributions map for the fin whale 20 Hz pulse and 130 Hz upsweep detected in the Northern Finnmark coastal region on 27 February 2014. The detected fin whale call rate spatial densities in units of number of calls per minute per 25 nmi^2 [(min) (5 nmi)^2] measured by POAWRS have peak values α indicated. The dashed contours represent the 50% PoD regions for the fin whale 20 Hz pulse (purple) and 130 Hz upsweep (green).

The fin whale 130 Hz upsweep vocalization source level estimates are found to be 20 dB lower than the source level estimates of their 20 Hz pulse vocalization (Figures 7 and 8). These results obtained here for the Norwegian Sea are consistent with a previous study that measured a difference of 24.5 ± 2.6 dB (Simon *et al.*, 2010) in received flux density levels between the fin whale 20 Hz pulse and 130 Hz upsweep in the Davis Strait after taking into account potential differences in transmission loss at 20 Hz and 130 Hz frequencies.

Due to the significantly lower source level of the fin whale 130 Hz upsweep vocalizations, we found that these signals were often undetectable in spectrograms produced from single omnidirectional hydrophone measurements. Beamforming with the coherent hydrophone array was necessary to enhance the SNR of the 130 Hz upsweep signal (by 13.7 dB in array gain) so that they were detectable above the ambient noise floor in the beamformed spectrograms (compare Figure 2a and c).

Advantages of using a large-aperture coherent hydrophone array to detect different types of fin whale vocalizations

An advantage of monitoring marine mammal vocalizations with a large-aperture coherent hydrophone array is that the vocalizations can be localized in bearing and range, and then mapped onto geographic space. The 50% PoD region for fin whale vocalizations extend over an area that is more than 200 km in diameter, enabling fin whales vocalizations distributed over wide areas to be simultaneously monitored, characterized, and localized using the coherent hydrophone array. The detected fin whale call rate spatial distribution maps shown here are useful for future studies of marine mammal behaviour as a function of concurrently measured environmental variates such as fish distributions or water-column temperature distributions. These future studies will provide insights into predator–prey dynamics occurring in the Norwegian Sea.

The fin whale 20 Hz pulse vocalizations have been the primary signal for passive acoustic monitoring of this marine mammal species due to the high source level and subsequently high SNR reception in single hydrophone or array sensor measurements. Furthermore, the 20 Hz pulse signal can travel long distances due to the signal's low water-column absorption losses, except for very shallow waters where modal cut-off occurs and there is penetration of signal energy into the sea bottom. The bearing-time estimates of fin whale 20 Hz pulses are typically noisier due to the poor angular resolution of most practical coherent hydrophone arrays at that frequency. The noise associated with the bearing estimates are problematic when many fin whales vocalize at multiple bearings that are in close proximity (see Figure 5b). Here, in addition to the 20 Hz pulse, the POAWRS system detected large volumes of fin whale 130 Hz upsweep. The high frequency 130 Hz upsweeps can be detected by the coherent hydrophone array with more accurate bearing-time estimates, because the array has better angular resolution (smaller bearing estimation error) at this frequency. The bearing-time trajectories of fin whale vocalizations at 130 Hz are much better resolved providing more accurate localization and geographic mapping of these vocalizations (see Figure 5b) compared with the 20 Hz pulse.

Coherent beamforming of the hydrophone array data is shown to significantly enhance the fin whale vocalization SNR and detection range by roughly two orders of magnitude over that of a single hydrophone (Figure 2). This implies that signals that are undetectable or barely audible on a single hydrophone can be pulled out of the limiting omnidirectional ambient noise floor with the coherent hydrophone array.

Insights into fin whale sound production

The mechanism for sound production in fin whales is still not well understood (Simon *et al.*, 2010). Previous measurements describe the fin whale 130 Hz upsweep calls as occurring in combination with their 20 Hz pulses (Clark and Gagnon, 2004; Simon *et al.*, 2010) leading to the question of whether the fin whale has control over production of the high frequency 130 Hz upsweep "or if they are an anatomically induced by-product from making

the 20 Hz pulse" (Simon *et al.*, 2010). During NorEx14, we found many fin whale vocalization bearing-time trajectories comprised of combinations of the 20 Hz pulse with the 130 Hz upsweep. We also found many fin whale vocalization bearing-time trajectories comprising solely of their 20 Hz pulses and also several of these comprised solely of their 130 Hz upsweeps. For instance, in the Lofoten region on 5 March 2014, we found a fin whale bearing-time trajectory where the fin whale 130 Hz upsweep was observed to occur in combination with the 20 Hz pulse for approximately 2.5 h, and then modified the vocalizations to just the 130 Hz upsweep for approximately 45 min. We localized the bearing-time trajectory of the fin whale 130 Hz upsweeps that did not occur in combination with the 20 Hz pulses, and found the fin whale located in an area with water depth greater than 100 m. This water depth is larger than the acoustic wavelength at 20 Hz and therefore favourable for acoustic propagation at this frequency. Furthermore, we estimated the PoD for the 20 Hz pulse at this fin whale location (which again takes into account the environmental parameters such as the range-dependent water depth and sound speed profiles) and estimated a 90% PoD value for the 20 Hz pulse. Given the water depth and the 90% PoD for the 20 Hz pulses, we concluded that we should have seen some 20 Hz pulses in a 45-min time interval if they had existed. The NorEx14 data set suggests that the fin whale does have control over production of the 130 Hz upsweep and can produce this vocalization type either solely or in combination with the 20 Hz pulse. We also find that the fin whale has control over production of the 18–19 Hz backbeat, since they can occur with or independently of the 20 Hz pulses.

The coherent hydrophone array measurement of the second-harmonic component of the fin whale primary 20 Hz pulse, detected at twice the peak centre frequency (approximately 41 Hz centred) and roughly 25–30 dB lower received pressure levels, provides insights into the mechanism for sound generation in the fin whale.

Conclusions

The vocalizations of the fin whale have been detected, characterized, and localized over wide areas of the Norwegian and Barents Seas based on observations from 18 February to 8 March 2014 using a large-aperture densely sampled coherent hydrophone array via the POAWRS technique. The received fin whale vocalizations in all regions observed are dominated by their characteristic 20 Hz pulses and high frequency 130 Hz upsweeps. An appreciable volume of fin whale large bandwidth 30–100 Hz downsweep chirp vocalizations were also received, as well as smaller amounts of their 18–19 Hz backbeat pulses. The time–frequency characteristics of these vocalization types and their occurrence rate time-series have been quantified for fin whales in three distinct regions of the Norwegian Sea, off the coasts of Alesund, Lofoten, and the Northern Finnmark. The detected fin whale diel vocalization rates are found to be significantly higher off the Northern Finnmark coast at roughly 37000 ± 5000 calls/day for the 20 Hz pulses and 21000 ± 5300 calls/day for the 130 Hz upsweeps. These detected call rates are a factor of roughly 5 times smaller for the 20 Hz pulses and roughly 17 times smaller for the 130 Hz upsweeps off the Alesund and Lofoten coasts. The detected fin whale vocalization rate spatial density distributions are mapped for their 20 Hz pulses and the 130 Hz upsweeps in all three observation regions of the Norwegian and Barents Seas. The vocalization source level distributions and PoD regions are estimated for the fin whale 20 Hz pulses and the 130 Hz upsweeps separately in the three distinct regions. The source levels in units of dB re 1 μPa at 1 m are found to be 190.5 ± 7.4 for the fin whale 20 Hz pulses and 170.3 ± 5.2 for the 130 Hz upsweeps. The fin whale 130 Hz upsweep vocalizations are received with significantly enhanced SNR by roughly 14 dB via beamforming with the coherent hydrophone array and are typically undetectable in single hydrophone measurements. For fin whales in close proximity to the coherent hydrophone array, the dominant and intense 20 Hz pulses were received along with their second-order harmonic, with a peak frequency centred at 41 Hz, and roughly 25–30 dB lower received pressure levels. Furthermore, from the large volumes of fin whale vocalizations observed, we find that fin whales are capable of producing each vocalization type either independently or simultaneously with their other call types. This study provides novel information on fin whale vocalization and sound production. Furthermore, the findings on fin whale vocalization distribution and behaviour can be applied to provide insights into predator–prey dynamics in important spawning areas along the Norwegian coast.

Acknowledgements

This research was supported by the US Office of Naval Research (Ocean Acoustics Program), the Norwegian Institute of Marine Research—Bergen, and the US National Science Foundation.

References

Andrews, M., Chen, T., and Ratilal, P. 2009. Empirical dependence of acoustic transmission scintillation statistics on bandwidth, frequency, and range in New Jersey continental shelf. The Journal of the Acoustical Society of America, 125: 111–124.

Andrews, M., Gong, Z., and Ratilal, P. 2011. Effects of multiple scattering, attenuation and dispersion in waveguide sensing of fish. The Journal of the Acoustical Society of America, 130: 1253–1271.

Baumgartner, M. F., and Mussoline, S. E. 2011. A generalized baleen whale call detection and classification system. The Journal of the Acoustical Society of America, 129: 2889–2902.

Becker, K., and Preston, J. 2003. The ONR five octave research array (FORA) at Penn State. *In* OCEANS 2003. Proceedings, vol. 5, pp. 2607–2610. IEEE.

Bergmann, P. G., Yaspan, A., Gerjuoy, E., Major, J., and Wildt, R. 1968. Physics of Sound in the Sea. Gordon and Breach, Philadelphia, PA.

Burdic, W. S. 1991. Underwater Acoustic System Analysis, pp. 322–360. Prentice Hall, New Jersey.

Castellote, M., Clark, C. W., and Lammers, M. O. 2012. Fin whale (*Balaenoptera physalus*) population identity in the western Mediterranean Sea. Marine Mammal Science, 28: 325–344.

Charif, R. A., Mellinger, D. K., Dunsmore, K. J., Fristrup, K. M., and Clark, C. W. 2002. Estimated source levels of fin whale (*Balaenoptera physalus*) vocalizations: adjustments for surface interference. Marine Mammal Science, 18: 81–98.

Clark, C., and Gagnon, G. 2004. Low-frequency vocal behaviors of baleen whales in the North Atlantic: insights from integrated undersea surveillance system detections, locations, and tracking from 1992 to 1996. Journal of Underwater Acoustics, 52: 48.

Clark, C. W., Borsani, J., and Notarbartolo-Di-sciara, G. 2002. Vocal activity of fin whales, *Balaenoptera physalus*, in the Ligurian Sea. Marine Mammal Science, 18: 286–295.

Clark, C. W., and Fristrup, K. M. 1997. Whales 95: a combined visual and acoustic survey of blue and fin whales off southern

California. Reports of the International Whaling Commission, 47: 583–600.

Clay, C. S., and Medwin, H. 1977. Acoustical oceanography: principles and applications. John Wiley & Sons, New York, NY. pp. 494–501.

Collins, M. D. 1993. A split-step padé solution for the parabolic equation method. The Journal of the Acoustical Society of America, 93: 1736–1742.

Croll, D. A., Clark, C. W., Acevedo, A., Tershy, B., Flores, S., Gedamke, J., and Urban, J. 2002. Bioacoustics: only male fin whales sing loud songs. Nature, 417: 809.

DiFranco, J., and Rubin, W. 1980. Radar Detection. Artech House Inc., Dedham, MA.

Edds, P. L. 1988. Characteristics of finback *Balaenoptera physalus* vocalizations in the St. Lawrence estuary. Bioacoustics, 1: 131–149.

Frieden, R. 2012. Probability, Statistical Optics, and Data Testing: A Problem Solving Approach, vol. 10. Springer Science & Business Media, Berlin, Germany.

Gaspà Rebull, O., Cusí, J. D., Ruiz Fernández, M., and Muset, J. G. 2006. Tracking fin whale calls offshore the Galicia Margin, North East Atlantic Ocean. The Journal of the Acoustical Society of America, 120: 2077–2085.

Gong, Z., Andrews, M., Jagannathan, S., Patel, R., Jech, J. M., Makris, N. C., and Ratilal, P. 2010. Low-frequency target strength and abundance of shoaling Atlantic herring (*Clupea harengus*) in the Gulf of Maine during the ocean acoustic waveguide remote sensing 2006 experiment. The Journal of the Acoustical Society of America, 127: 104–123.

Gong, Z., Jain, A. D., Tran, D., Yi, D. H., Wu, F., Zorn, A., Ratilal, P., *et al.* 2014. Ecosystem scale acoustic sensing reveals humpback whale behavior synchronous with herring spawning processes and re-evaluation finds no effect of sonar on humpback song occurrence in the Gulf of Maine in fall 2006. PLoS One, 9: e104733.

Gong, Z., Ratilal, P., and Makris, N. C. 2015. Simultaneous localization of multiple broadband non-impulsive acoustic sources in an ocean waveguide using the array invariant. The Journal of the Acoustical Society of America, 138: 2649–2667.

Gong, Z., Tran, D. D., and Ratilal, P. 2013. Comparing passive source localization and tracking approaches with a towed horizontal receiver array in an ocean waveguide. The Journal of the Acoustical Society of America, 134: 3705–3720.

Harris, D., Matias, L., Thomas, L., Harwood, J., and Geissler, W. H. 2013. Applying distance sampling to fin whale calls recorded by single seismic instruments in the northeast Atlantic. The Journal of the Acoustical Society of America, 134: 3522–3535.

Hirose, K., Kawano, S., Konishi, S., and Ichikawa, M. 2011. Bayesian information criterion and selection of the number of factors in factor analysis models. Journal of Data Science, 9: 243–259.

Huang, W. 2017. Temporal-spectral characterization and classification of marine mammal vocalizations and diesel-electric ships radiated sound over continental shelf scale regions with coherent hydrophone array measurements. PhD thesis, Northeastern University, MA, USA.

Huang, W., Wang, D., Garcia, H., Godø, O. R., and Ratilal, P. 2017. Continental shelf-scale passive acoustic detection and characterization of diesel-electric ships using a coherent hydrophone array. Remote Sensing, 9: 772.

Huang, W., Wang, D., and Ratilal, P. 2016. Diel and spatial dependence of humpback song and non-song vocalizations in fish spawning ground. Remote Sensing, 8: 712.

Jagannathan, S., Bertsatos, I., Symonds, D., Chen, T., Nia, H. T., Jain, A. D., Andrews, M., *et al.* 2009. Ocean acoustic waveguide remote sensing (OAWRS) of marine ecosystems. Marine Ecology Progress Series, 395: 137–160.

Jain, A. D. 2015. Instantaneous continental-shelf scale sensing of cod with Ocean Acoustic Waveguide Remote Sensing (OAWRS). PhD thesis, Doctoral dissertation, Massachusetts Institute of Technology, MA, USA.

Jensen, F., Kuperman, W., Porter, M., and Schmidt, H. 2011. Computational Ocean Acoustics, pp. 708–713. Springer Science & Business Media, Berlin, Germany.

Johnson, D. H., and Dudgeon, D. E. 1992. Array Signal Processing: Concepts and Techniques. Simon & Schuster, New York, NY.

Jolliffe, I. 2002. Principal Component Analysis, 2nd edn., pp. 11–31. Wiley Online Library, New York.

Kanungo, T., Mount, D. M., Netanyahu, N. S., Piatko, C. D., Silverman, R., and Wu, A. Y. 2002. An efficient k-means clustering algorithm: analysis and implementation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 24: 881–892.

Kay, S. M. 1998. Fundamentals of Statistical Signal Processing, Vol. II: Detection Theory. Prentice Hall, Upper Saddle River, NJ.

Kinsler, L. E., Frey, A. R., Coppens, A. B., and Sanders, J. V. 1999. Fundamentals of acoustics. Fundamentals of Acoustics, 4th edn. Ed. by L. E. Kinsler, A. R. Frey, A. B. Coppens, and J. V. Sanders. Wiley-VCH, New York, NY. 560 pp.

Klinck, H., Nieukirk, S. L., Mellinger, D. K., Klinck, K., Matsumoto, H., and Dziak, R. P. 2012. Seasonal presence of cetaceans and ambient noise levels in polar waters of the north Atlantic. The Journal of the Acoustical Society of America, 132: EL176–EL181.

Madsen, P. T., and Wahlberg, M. 2007. Recording and quantification of ultrasonic echolocation clicks from free-ranging toothed whales. Deep Sea Research Part I: Oceanographic Research Papers, 54: 1421–1444.

Makris, N. C. 1996. The effect of saturated transmission scintillation on ocean acoustic intensity measurements. The Journal of the Acoustical Society of America, 100: 769–783.

Makris, N. C., Avelino, L. Z., and Menis, R. 1995. Deterministic reverberation from ocean ridges. The Journal of the Acoustical Society of America, 97: 3547–3574.

Makris, N. C., Ratilal, P., Jagannathan, S., Gong, Z., Andrews, M., Bertsatos, I., Godø, O. R., *et al.* 2009. Critical population density triggers rapid formation of vast oceanic fish shoals. Science, 323: 1734–1737.

Makris, N. C., Ratilal, P., Symonds, D. T., Jagannathan, S., Lee, S., and Nero, R. W. 2006. Fish population and behavior revealed by instantaneous continental shelf-scale imaging. Science, 311: 660–663.

Matias, L., and Harris, D. 2015. A single-station method for the detection, classification and location of fin whale calls using ocean-bottom seismic stations. The Journal of the Acoustical Society of America, 138: 504–520.

McDonald, M. A., Hildebrand, J. A., and Webb, S. C. 1995. Blue and fin whales observed on a seafloor array in the northeast pacific. The Journal of the Acoustical Society of America, 98: 712–721.

Nakken, O. 2008. Norwegian Spring-Spawning Herring & Northeast Arctic Cod: 100 Years of Research Management. Tapir Academic Press, Trondheim, Norway.

Nieukirk, S. L., Stafford, K. M., Mellinger, D. K., Dziak, R. P., and Fox, C. G. 2004. Low-frequency whale and seismic airgun sounds recorded in the mid-Atlantic Ocean. The Journal of the Acoustical Society of America, 115: 1832–1843.

Northrop, J., Cummings, W., and Thompson, P. 1968. 20-Hz signals observed in the central pacific. The Journal of the Acoustical Society of America, 43: 383–384.

Ratilal, P., Lai, Y., Symonds, D. T., Ruhlmann, L. A., Preston, J. R., Scheer, E. K., Garr, M. T., *et al.* 2005. Long range acoustic imaging of the continental shelf environment: the acoustic clutter reconnaissance experiment 2001. The Journal of the Acoustical Society of America, 117: 1977–1998.

Richard, O. D., Peter, E. H., and David, G. S. 2001. Pattern Classification. A Wiley-Interscience, New York, NY. pp. 373–378.

Schory, G. 2015. Statistical analysis of acoustic transmission scintillation in the 2014 Nordic Seas Experiment. Master's thesis, Massachusetts Institute of Technology, MA, USA.

Sezan, M. I. 1990. A peak detection algorithm and its application to histogram-based image data reduction. Computer Vision, Graphics, and Image Processing, 49: 36–51.

Shapiro, A. D., and Wang, C. 2009. A versatile pitch tracking algorithm: from human speech to killer whale vocalizations. The Journal of the Acoustical Society of America, 126: 451–459.

Simon, M., Stafford, K. M., Beedholm, K., Lee, C. M., and Madsen, P. T. 2010. Singing behavior of fin whales in the Davis Strait with implications for mating, migration and foraging. The Journal of the Acoustical Society of America, 128: 3200–3210.

Širović, A., Hildebrand, J. A., and Wiggins, S. M. 2007. Blue and fin whale call source levels and propagation range in the southern ocean. The Journal of the Acoustical Society of America, 122: 1208–1215.

Širović, A., Williams, L. N., Kerosky, S. M., Wiggins, S. M., and Hildebrand, J. A. 2013. Temporal separation of two fin whale call types across the eastern north pacific. Marine Biology, 160: 47–57.

Stimpert, A. K., DeRuiter, S. L., Falcone, E. A., Joseph, J., Douglas, A. B., Moretti, D. J., Friedlaender, A. S., *et al.* 2015. Sound production and associated behavior of tagged fin whales (*Balaenoptera physalus*) in the Southern California Bight. Animal Biotelemetry, 3: 23.

Thompson, P. O., Findley, L. T., and Vidal, O. 1992. 20-hz pulses and other vocalizations of fin whales, *Balaenoptera physalus*, in the Gulf of California, Mexico. The Journal of the Acoustical Society of America, 92: 3051–3057.

Tran, D., Andrews, M., and Ratilal, P. 2012. Probability distribution for energy of saturated broadband ocean acoustic transmission: results from Gulf of Maine 2006 experiment. The Journal of the Acoustical Society of America, 132: 3659–3672.

Tran, D. D., Huang, W., Bohn, A. C., Wang, D., Gong, Z., Makris, N. C., and Ratilal, P. 2014. Using a coherent hydrophone array for observing sperm whale range, classification, and shallow-water dive profiles. The Journal of the Acoustical Society of America, 135: 3352–3363.

Urick, R. J. 1983. Principles of underwater sound, 3rd ed. McGraw Hill, New York, NY. pp. 29–65, 328–366.

Wang, C., and Seneff, S. 2000. Robust pitch tracking for prosodic modeling in telephone speech. *In* 2000 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2000. ICASSP'00, Proceedings, vol. 3, pp. 1343–1346. IEEE, Istanbul, Turkey.

Wang, D., Garcia, H., Huang, W., Tran, D. D., Jain, A. D., Yi, D. H., Gong, Z., *et al.* 2016. Vast assembly of vocal marine mammals from diverse species on fish spawning ground. Nature, 531: 366–370.

Wang, D., Huang, W., Garcia, H., and Ratilal, P. 2016. Vocalization source level distributions and pulse compression gains of diverse baleen whale species in the Gulf of Maine. Remote Sensing, 8: 881.

Wang, D., and Ratilal, P. 2017. Angular resolution enhancement provided by nonuniformly-spaced linear hydrophone arrays in ocean acoustic waveguide remote sensing. Remote Sensing, 9: 1036.

Watkins, W. A. 1981. Activities and underwater sounds of fin whales. The Scientific Reports of the Whales Research Institute, 33: 83–117.

Watkins, W. A., Tyack, P., Moore, K. E., and Bird, J. E. 1987. The 20-Hz signals of finback whales (*Balaenoptera physalus*). The Journal of the Acoustical Society of America, 82: 1901–1912.

Weirathmueller, M. J., Wilcock, W. S., and Soule, D. C. 2013. Source levels of fin whale 20 Hz pulses measured in the Northeast Pacific Ocean. The Journal of the Acoustical Society of America, 133: 741–749.

Appendix 1.

Modelling probability of detection regions for fin whale vocalizations

Here the approach for calculating the POAWRS PoD $P_{\mathrm{D}}(r)$ for a fin whale vocalization as a function of range r from the coherent hydrophone array is described. For a fin whale at range r from the POAWRS receiver array, its vocalization signal can be detected above the ambient noise if the sonar equation (Clay and Medwin, 1977; Urick, 1983; Burdic, 1991; Jensen *et al.*, 2011; Gong *et al.*, 2014) is satisfied:

$$\mathrm{NL} + \mathrm{DT} - \mathrm{AG} < L_{\mathrm{S}} - \mathrm{TL}(r), \tag{A.1}$$

where L_{S} is the source level of the fin whale vocalization, NL is the ambient noise level in the frequency band of the fin whale vocalization signal, AG is the coherent beamforming gain of the passive receiver array, DT is the detection threshold, and TL is the broadband transmission loss.

The fin whale vocalization signals are detected from the beamformed spectrograms and typically occupy roughly M number of independent time–frequency pixels $\Delta f \Delta t$. We first calculate the detection probability $p_{\mathrm{D},1}(r)$ in a single frequency–time pixel using (DiFranco and Rubin, 1980; Urick, 1983)

$$p_{\mathrm{D},1}(r) = \int_{-\infty}^{+\infty} f_{L_{\mathrm{R}}}(L_{\mathrm{R}}(r)) \int_{-\infty}^{L_{\mathrm{R}}(r)-\mathrm{DT}} f_{L_{\mathrm{N}}}(L_{\mathrm{N}}) \mathrm{d}L_{\mathrm{N}} \mathrm{d}L_{\mathrm{R}}, \tag{A.2}$$

where $f_{L_{\mathrm{N}}}(L_{\mathrm{N}})$ is the probability density function of the log-transformed ambient noise pressure-squared $L_{\mathrm{N}}(t,f) = 10\log_{10}(|P_{\mathrm{N}}(t,f)/P_{\mathrm{ref}}|^2) = S_{\mathrm{N}}(t,f) + 10\log_{10}(\Delta f) - \mathrm{AG}$ within a single beamformed spectrogram time–frequency pixel in the frequency range of the fin whale vocalization, where $P_{\mathrm{N}}(t,f)$ is the ambient noise pressure at time t within frequency bin Δf centred at frequency f and $S_{\mathrm{N}}(t,f)$ is the omnidirectional ambient noise spectral density level; $f_{L_{\mathrm{R}}}(L_{\mathrm{R}}(r))$ is the probability density function of the received fin whale vocalization signal log-transformed pressure-squared $L_{\mathrm{R}}(r|t,f) = 10\log_{10}(|P_{\mathrm{R}}(r|t,f)/P_{\mathrm{ref}}|^2) = L_{\mathrm{S}} - \mathrm{TL}(r) + 10\log_{10}\frac{\Delta f}{B(t)}$ within a single beamformed spectrogram time–frequency pixel, where $P_{\mathrm{R}}(r|t,f)$ is the received fin whale vocalization signal pressure, and $B(t)$ is the instantaneous bandwidth of that signal at time t. The number of independent beamformed spectrogram frequency–time pixels occupied by the fin whale vocalization signal is related to the instantaneous bandwidth via $M\Delta f\Delta t = \tau B(t)$, where τ is the signal duration. An exponential Gamma distribution (Bergmann *et al.*, 1968; Makris, 1996; Tran *et al.*, 2012) describes the log-transformed ambient noise pressure-squared and log-transformed received fin whale vocalization pressure-squared within a single beamformed spectrogram time-frequency pixel:

$$f_{L_{\mathrm{N}}}(L_{\mathrm{N}}) = \frac{1}{(10\log_{10}e)\Gamma(\mu)} \left(\frac{\mu}{\langle P_{\mathrm{N}}^2\rangle}\right)^{\mu} 10^{\mu L_{\mathrm{N}}/10} \exp\left(-\mu\frac{10^{L_{\mathrm{N}}/10}}{\langle P_{\mathrm{N}}^2\rangle}\right) \tag{A.3}$$

$$f_{L_R}(L_R(r)) = \frac{1}{(10\log_{10}e)\Gamma(\mu)}\left(\frac{\mu}{\langle P_R^2(r)\rangle}\right)^{\mu} 10^{\mu L_R(r)/10} \exp\left(-\mu\frac{10^{L_R(r)/10}}{\langle P_R^2(r)\rangle}\right), \tag{A.4}$$

where μ is the time-bandwidth product or number of statistically independent fluctuations of the respective pressure-squared quantities. Since the beamformed spectrograms have time–frequency pixels that satisfy $\Delta f \Delta t = 1$, both the ambient noise level and the received fin whale vocalization signal level within each beamformed spectrogram time–frequency pixel can be treated as instantaneous with time-bandwidth product $\mu = 1$ and 5.6 dB standard deviation. For the received fin whale vocalization signal level, this standard deviation includes both the standard deviation of the fin whale vocalization source level, as well as the standard deviation of the broadband waveguide transmission loss. The 5.6 dB standard deviation used here for the received fin whale vocalization signal level is a good approximation to the standard deviations shown in Figures 7 and 8.

We assume that the received fin whale vocalization signal is detectable if it stands above the ambient noise in at least 30% of the M time–frequency pixels of the beamformed spectrogram. The overall PoD, $P_D(r)$, for the fin whale vocalization signal as a function of range r from the POAWRS receiver array is then calculated from the Gaussian approximation to the binomial cumulative distribution function (CDF) (DiFranco and Rubin, 1980) as:

$$P_D(r) = 1 - \Phi\left(\frac{0.3M - Mp_{D,1}(r)}{\sqrt{Mp_{D,1}(r)(1-p_{D,1}(r))}}\right). \tag{A.5}$$

where $\Phi(z) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z} e^{-u^2/2}\mathrm{d}u$. The Gaussian approximation to the binomial CDF is an appropriate model for the overall performance of the detector when considering the thousands of fin whale vocalization signals analysed. The exponential Gamma distribution (Makris, 1996; Tran *et al.*, 2012) for the log-transform of Gaussian field measurements, used here to model the probability density function of the received fin whale vocalization signal level and the ambient noise level, has been calibrated with thousands of log-transformed intensity measurements from controlled source transmissions made during two past experiments conducted in a similar continental shelf environment (Andrews *et al.*, 2009; Tran *et al.*, 2012), as well as in the Norwegian Sea (Schory, 2015).

The fin whale vocalization source levels L_S used here are estimated from a subset of the POAWRS received vocalizations (see Figures 7 and 8), in units of dB re 1 μPa at 1m, and are provided in Table 5. The omnidirectional ambient noise spectral density levels $S_N(f)$ are estimated directly from the POAWRS receiver array using data segments that are devoid of fin whale vocalizations and other acoustic sound sources. The $S_N(f)$ used are provide in Table 5.

This approach was previously formulated and applied to model the POAWRS PoD regions for vocalizations from diverse marine mammal species in the Gulf of Maine, that include fin whale vocalizations, received by a coherent hydrophone array (see Section I of the Supplementary Information of Wang *et al.*, 2016a). The approach has also been extended and applied to model POAWRS PoD regions for underwater sound radiated by diesel-electric ships and received on a coherent hydrophone array (Huang *et al.*, 2017; Huang, 2017).

Handling editor: Francis Juanes